

Etäyhteydellä toteutetun henkilöarvioinnin yhteys työmenestykseen

Psykologian pro gradu -tutkielma

Laatija:

Matias Kaasalainen

Ohjaajat:

Psykologian professori, Turun Yliopisto, Paula Salo

Tutkimusjohtaja, Psycon Oy, FT (psykologia), Mikael Nederström

06.10.2021

Turku

Turun yliopiston laatu järjestelmän mukaisesti tämän julkaisun alkuperäisyys on tarkastettu
Turnitin OriginalityCheck -järjestelmällä.

Pro gradu -tutkielma

Oppiaine: Psykologia

Tekijä(t): Matias Kaasalainen

Otsikko: Etäyhteydellä toteutetun henkilöarvioinnin yhteys työmenestykseen

Ohjaaja(t): Professori Paula Salo, Tutkimusjohtaja Mikael Nederström

Sivumäärä: 33 sivua

Päivämäärä: 06.10.2021

Viimeistään koronapandemian myötä henkilöarviointien toteuttaminen etäyhteydellä on lisääntynyt voimakkaasti, vaikka muutos kohti digitaalisuutta ja paikkariippumattomuutta on ollut käynnissä jo pidemmän aikaa. Aikaisempi tutkimus etäyhteydellä toteutetusta henkilöarvioinnista on tyypillisesti keskittynyt esimerkiksi prosessin pelillistämiseen tai tekoälyn hyödyntämiseen arvioinnissa, mutta perinteisemmille psykologiselle arviointimenetelmille pohjautuvasta henkilöarvioinnista etäyhteydellä on olemassa niukasti tietoa. Tässä tutkimuksessa tarkasteltiin, kuinka hyvin etäyhteydellä toteutettu henkilöarviointi ennustaa esihenkilön arvioimaa rekrytoidun henkilön työmenestystä verrattuna perinteiseen henkilöarviointiin. Lisäksi tutkittiin, mitkä psykologiset ominaisuudet ovat vahvimmin yhteydessä erinomaiseen työmenestykseen. Tutkimus tuotti uutta tietoa, minkä lisäksi henkilöarviointien laadun seurantatutkimus on osa hyvää henkilöarviointikäytäntöä. Tutkimuksen aineisto koostui esihenkilöille suunnatun seurantatutkimuksen tuloksista sekä Psycon Oy:n henkilöarviointiaineistosta. Rekrytoidut henkilöt olivat työskennelleet uudessa työssään vähintään puoli vuotta työmenestyksen mittaushetkellä. Henkilöarviointien ennustavuutta tutkittiin vertaamalla rekrytoitujen henkilöiden henkilöarviointituloksia esihenkilön arvioimaan työmenestykseen. Otokoko oli etätilanteessa 96 ja läsnätilanteessa 90 eli yhteensä 186. Tutkimuksessa havaittiin, että etäyhteydellä toteutettu henkilöarviointi ja perinteinen henkilöarviointi ennustivat työmenestystä yhtä hyvin. Yksittäiset ominaisuudet olivat melko heikosti yhteydessä työmenestykseen. Tässä tutkimuksessa numeerinen ja induktiivinen päättelykyky ennustivat erinomaista työmenestystä, kun henkilön estyneisyys ja defensiivisyys puolestaan heikensivät työmenestyksen ennustetta. Tämän tutkimuksen tulokset ovat lupaavia psykologisen etäarvioinnin toimivuuden suhteen, mutta vaihteluvälin kapeutuminen aineistossa hankaloittaa tulosten yleistettävyyttä. Jatkossa etähenkilöarvioinnin ennustavuutta on hyvä tutkia myös koronapandemia-tilanteen helpottuttua, jolloin työnkuvat eivät todennäköisesti rajoitu yhtä vahvasti pelkkään etätyöskentelyyn. Lisäksi on syytä kiinnittää erityistä huomiota käytettyihin työmenestyksen mittareihin sekä tutkimuksen otantaan.

Avainsanat: henkilöarviointi, ennustavuus, etätyö, työmenestys

Sisällysluettelo

1	Johdanto	1
1.1	Psykologinen henkilöarviointi	2
1.2	Henkilöarvioinnin ennustavuus aiemmissa tutkimuksissa	3
1.3	Psykologinen henkilöarviointi etäyhteydellä	5
1.3.1	Sähköisten arviointimenetelmien edut	5
1.3.2	Sähköisiin arviointimenetelmiin liittyvät ongelmakohdat	7
1.3.3	Vuorovaikutukselliset menetelmät psykologisessa etäarvioinnissa	8
1.4	Tutkimusongelmat ja hypoteesit	10
2	Menetelmät	11
2.1	Tutkimusasetelma ja tutkittavat	11
2.2	Henkilöarviointimenetelmät	12
2.2.1	Kognitiiviset kykytehtävät	13
2.2.2	Motivaatiomittari	14
2.2.3	Stressireaktiokysely	15
2.3	Toteutunut työmenestys	16
2.4	Analyysit ja kriteerit	17
3	Tulokset	19
3.1	Henkilöarvioinnin yleisarvion ennustavuus	19
3.2	Teorialähtöinen logistinen regressioanalyysi	21
3.3	Aineistolähtöinen logistinen regressioanalyysi	22
4	Pohdinta	24
4.1	Henkilöarvioinnin osuvuus	24
4.2	Yksittäisten menetelmien ennustavuus	27
4.3	Tutkimuksen vahvuudet ja rajoitteet	29
4.4	Jatkotutkimusehdotukset	32
	Lähteet:	34

1 Johdanto

Koronapandemian myötä henkilöarviointeja on viimeisen vuoden ajan toteutettu aiemmasta poiketen ensisijaisesti etäyhteydellä ja jo ennen pandemiaa digitaaliset ratkaisut ovat kasvattaneet merkittävästi rooliaan työelämässä. Tämän vuoksi on tärkeää tutkia, kuinka hyvin etäyhteydellä toteutettu henkilöarviointi ennustaa todellista työmenestystä. Lisäksi tarvitaan lisää tietoa siitä, eroavatko lähikontaktissa toteutettu henkilöarviointi ja etäyhteydellä toteutettu henkilöarviointi toisistaan työmenestyksen ennustavuuden suhteen, sillä tämänkaltaista vertailevaa tutkimusta ei ole aikaisemmin tehty. Tyypillisesti työmenestyksen mittarina käytetään esihenkilön tekemää arviota työhön valitusta henkilöstä (Sackett ym., 2017), mutta myös muut mittarit, kuten palkkakehityksen seuranta, eteneminen ja vertaisarviot ovat mahdollisia (Bartels ym., 2000; Damitz ym., 2003; Jansen & Stoop, 2001).

Organisaatiot hyödyntävät henkilöarviointeja tavallisesti työhönoton yhteydessä minimoidakseen rekrytointiin liittyvät riskit ja lisätäkseen varmuutta siitä, että valittu henkilö tulee pärjäämään työssä (Honkanen, 2005). Väärälle rekrytoinnille voi arvioiden mukaan kertyä hintaa jopa kolmannes henkilön vuosipalkasta (Montgomery-Young, 2018) ja lisäksi epäsuoria kustannuksia kuten projektien keskeytyksiä (Karsan, 2007), mikä osaltaan korostaa osuvan henkilöarvioinnin merkitystä osana rekrytointipäätöstä. Henkilöarviointien käyttö Suomessa onkin verrattain yleistä ja yhden arvion mukaan Suomessa tehdään vuosittain 10 000–15 000 henkilöarviointia rekrytointien yhteydessä (Puttonen, 2016).

Etätyösuositusten myötä aikaisemmin toimistoympäristössä sovellettuja henkilöarviointimenetelmiä on toteutettu etäyhteydellä ennennäkemättömässä mittakaavassa. Vaikka esimerkiksi psykologisten arviointimenetelmien sähköistäminen on ollut käynnissä jo pitkään, herää kysymys siitä, onko yllättävän pandemian aiheuttama äkillinen pakkosiirtymä etätyöskentelyyn vaikuttanut henkilöarviointien laatuun. Aikaisemmat tutkimukset etäyhteydellä toteutettavasta psykologisesta arvioinnista ovat tuottaneet sekä myönteisiä että kielteisiä tuloksia (Barak & Buchanan, 2004; Barak & English, 2002; Blacksmith ym., 2016; Campos ym., 2011; Meade ym., 2007; Rutherford ym., 2016; Sampson, 2000; Tippins ym., 2006; Wall, 2000).

Toistaiseksi etähenkilöarviointia koskeva tutkimus ei kuitenkaan vielä kata aukottomasti kaikkia teemaan liittyviä kokonaisuuksia, minkä vuoksi aiheetta on syytä tutkia lisää. Lisää selvyyttä vaaditaan esimerkiksi siihen, miten etähenkilöarviointi toimii kokonaisuutena

suhteessa perinteiseen henkilöarviointiin, jossa vähintään osa arvioinnista suoritetaan lähikontaktissa toimistolla. Tutkimukset etähenkilöarvioinnista ovat usein keskittyneet tekoälyn tai pelillistämisen hyödyntämiseen esimerkiksi persoonallisuuden arvioimisessa (Howland ym., 2015; Kosinski ym. 2013; Youyou ym., 2015), mutta niin sanottuihin perinteisiin henkilöarviointimenetelmiin, kuten haastatteluun, kognitiivisiin kykytehtäviin, kyselyihin ja simulaatioihin pohjautuvan henkilöarvioinnin tutkimus etäkontekstissa on jäänyt liian vähälle huomiolle. Näin ollen etähenkilöarvioinnin toimivuuden arvioiminen tutkimuskirjallisuuden perusteella pohjautuu tässä kohtaa pitkälti yksittäisistä osa-alueista, kuten kognitiivisista kykytehtävistä, kyselymenetelmistä ja haastatteluista etäkontekstissa tehtyihin tutkimuksiin. Kun huomioon otetaan kommunikaatioteknologioiden yleisesti kasvanut merkitys työelämässä ja koronapandemian aiheuttama, todennäköisesti pysyvä muutos etätöön ja paikkariippumattomuuden yleistymisessä, voidaan psykologisen etähenkilöarvioinnin laatututkimusta pitää kriittisenä alan tulevaisuuden kannalta. Lisäksi henkilöarvioinnin seurantatutkimus on keskeinen osa hyvää henkilöarviointikäytäntöä (Suomen Psykologiliitto ry, 2019).

Tässä tutkimuksessa tarkastelen, kuinka hyvin etäyhteydellä toteutettu henkilöarviointi Psycon Oy:llä ennustaa esihenkilön arvioimaa rekrytoidun työntekijän todellista työmenestystä. Lisäksi tutkin, kuinka hyvin etähenkilöarviointi toimii verrattuna niin sanottuun perinteiseen henkilöarviointiin sekä sitä, mitkä kognitiivisten kykytehtävien ja eri kyselymenetelmien dimensioista ennustavat työmenestystä parhaiten. Tutkimus on osa Psycon Oy:n toteuttamaa seurantatutkimusta.

1.1 Psykologinen henkilöarviointi

Tyypillisesti psykologisen henkilöarvioinnin tarkoituksena on tarkastella yksilön soveltuvuutta tiettyyn työtehtävään ja työympäristöön sekä tarjota työnantajalle tietoa päätöksenteon tueksi (Suomen Psykologiliitto ry, 2019). Rekrytointien lisäksi henkilöarviointeja voidaan hyödyntää myös muissa yhteyksissä, kuten urasuunnittelussa tai osaamisen arvioimisen ja kehittämisen tukena (Honkanen, 2005). Henkilöarvioinnin tulee perustua erikseen määriteltyihin kriteereihin, jotka työnantaja on käynyt läpi yhdessä arvioijan kanssa. Hyvässä henkilöarvioinnissa hakijaa arvioidaan suhteessa näihin kriteereihin, menetelmät ovat luotettavia ja arvioinnin suorittavat alan sertifioidut ammattilaiset. Henkilöarvioinnin tulokset raportoidaan kirjallisessa ja/tai suullisessa muodossa ja palaute annetaan sekä organisaatiolle että arvioitavalle itselleen. Raporttiin sisältyy tyypillisesti sekä numeerisia että kirjallisia

arvioita arvioiduista osa-alueista, kuten myös numeerinen yleisarvio henkilön soveltuvuudesta tehtävään (overall assessment rating; OAR) (Hermelin ym., 2007; Honkanen, 2005). Yleisarviossa otetaan huomioon niin työnhakijan vahvuudet kuin kehityskohteet suhteessa työtehtävälle asetettuihin kriteereihin.

Psykologisen henkilöarvioinnin tutkimuksen kohteena ovat muun muassa yksilön kognitiivinen kapasiteetti, työskentelytyyli, ajattelu- ja päätöksentekotapa, yhteistyötaidot ja sosiaalinen tyyli sekä persoonallisuus (Honkanen, 2005; Suomen Psykologiliitto ry, 2019). Psycon Oy:lla henkilöarviointi sisältää käytännössä aina kognitiivisen kyvykkyyden mittarit, erilaisia persoonallisuutta ja käyttäytymistä kartoittavia kyselymenetelmiä eli inventaareja, konsultin puolistrukturoidun haastattelun, ryhmäsimulaation sekä mahdollisia yksilösimulaatioita, joissa arvioidaan esimerkiksi henkilön päätöksentekotapaa tai esiintymistaitoja. Lisäksi henkilöarviointiassistentit tuottavat havaintoja arvioitavan käyttäytymisestä. Tarkoituksena on arvioida hakijan soveltuvuutta monimenetelmällisesti, sillä useiden rinnakkaisten mittareiden on havaittu parantavan henkilöarvioinnin laatua (Woehr & Arthur, 2003).

1.2 Henkilöarvioinnin ennustavuus aiemmissä tutkimuksissa

Henkilöarvioinnin ennustavuustutkimusta on tehty paljon, mutta tutkimustuloksia tarkasteltaessa on syytä huomioda aika, jolloin tutkimus on tehty. Alan tutkimus on osittain hyvin vanhaa ja monen tuoreemmankin meta-analyysin tutkimusaineisto on yleensä monen vuoden takaa. Tämä on pulmallista jo siitä syystä, että arviointikäytännöt ovat voineet muuttua ajan myötä. Kun huomioidaan lisäksi sekä arviointityön että koko työelämän nopea muutos 2010-luvulla esimerkiksi digitalisaation ja työn uusien muotojen vuoksi, on vanhojen tutkimusten tuloksiin suhtauduttava kriittisesti.

Henkilöarvioinnilla on useiden tutkimusten perusteella ennustevoimaa (Arthur ym., 2003; Dayan ym., 2002; Heikkilä, 2017; Hermelin ym., 2007; Klimoski & Brickner, 1987; Lievens ym., 2003; Morris ym., 2015; Sackett ym., 2017; Schmidt & Hunter, 1998; Schmitt ym., 1984). Tyypillisimmin henkilöarvioinnin ennustavuutta on tutkittu tarkastelemalla henkilöarvioinnin yleisarvion (OAR) yhteyttä esihenkilöiden tekemiin arvioihin alaisten työsuoritusarvioista (Sackett ym., 2017). Esihenkilöarviot voivat sisältää kokonaisarvion työsuoritusarvioista, kuten myös kommunikaatiotaitojen, yhteistyötaitojen, sopeutuvuuden, ongelmanratkaisukyvyyn, työhön liittyvien tietojen tai muiden yksittäisten aladimensioiden arviointia (Lievens ym., 2003; Sackett ym., 2017). Myös kollegojen arvioita, opintomenestystä sekä palkka- ja urakehitystä tarkastelleiden tutkimusten perusteella henkilöarvioinnilla voidaan ennakoita työsuoritusarvioita

(Bartels ym., 2000; Damitz ym., 2003; Jansen & Stoop, 2001). Meta-analyysissä esihenkilöarvioiden arvioijien väliseksi (inter-rater reliability) reliabiliteetiksi on arvioitu .52 ja saman arvioijan antamien arvioiden (intra-rater reliability) reliabiliteetiksi .86 (Viswesvaran, 1996). Tutkimuksen piirissä sitä, että henkilöarvioinnin tuloksia verrataan suhteessa edellä esitetyn kaltaisiin suoriutumisen kriteereihin, voidaan kutsua kriteerivaliditeettitutkimukseksi. Sitä kuinka hyvin henkilöarviointi kykenee ennustamaan työsuoriutumista näillä kriteereillä mitattuna, puolestaan kutsutaan ennustevaliditeetiksi.

Huomioitavaa on, että tutkimustuloksissa esiintyy myös jonkin verran hajontaa. Meta-analyysien vaihtelevia tuloksia on selitetty mukaan otettujen henkilöarviointien vaihtelevalla laadulla, metodologisilla tekijöillä, kuten virheellisten reliabiliteettiestimaattien käytöllä mittausvirhekorjauksissa (Viswesvaran ym., 1996) sekä liian vähäisellä vaihtelulla hakijoiden kompetensseissa, mikä johtaa vaihteluvälin kapeutumiseen (Lievens & Thornton, 2005). Vaihteluvälin kapeutumisella tarkoitetaan tilannetta, jossa tutkimusaineiston varianssi ei vastaa varianssia kohdepopulaatiossa (Dahlke & Wiernik, 2020). Henkilöarviointien seuranta tutkimuksessa tutkimus kohdistuu tyypillisesti vain arvioinnissa parhaiten menestyneisiin ja sen myötä valituiksi tulleeisiin, eikä kaikkiin henkilöarviointiin osallistuneisiin henkilöihin. Työhön rekrytoidut henkilöt ovat toisin sanoen niin valikoitunut joukko, ettei aineistossa esiinny tilastotieteen näkökulmasta riittävästi varianssia. Hermelin ja kumppanit (2007) argumentoivat, että vaihteluvälin kapeutuminen näkyy myös matalampina validiteettiestimaatteina, eikä ongelmasta päästä aina pelkällä tilastollisella korjaamisella.

Tutkimuksenkin piirissä paljon käytetty henkilöarvioinnin yleisarvio (OAR) on monin tavoin käytännöllinen mittari ja sen etuna voidaan pitää sitä, että se perustuu monelle eri käyttäytymisen mittarille (Jansen & Stoop, 2001). Toisaalta pelkän yleisarvion voidaan ajatella yksinkertaistavan liikaa yksilön soveltuvuusrakennetta. Joissakin tapauksissa yksittäisten aladimensioiden avulla työssä pärjäämistä on pystytty ennustamaan jonkin verran pelkkää yleisarviota paremmin (Arthur ym., 2003; Meriac ym., 2008). Tästä syystä on tärkeää, että henkilöarviointitulosten yhteydessä raportoidaan yleisarvion lisäksi myös arviot niistä tarkkarajaisemmista osa-alueista, jotka ovat kiinnostuksen kohteena. Arthurin ja kumppanien (2003) tutkimuksessa parhaiten työssä pärjäämistä ennustaviksi kyvyiksi nousivat kyky huomioda muut, kommunikointitaidot, tarmokkuus, kyky vaikuttaa muihin sekä organisointi- ja ongelmanratkaisukyky. Lienee kuitenkin perustelua todeta, että aladimensioiden oleellisuus on aina suhteessa myös kyseessä olevan tehtävän vaatimuksiin. Lisäksi tärkeä kysymys on se, miten henkilöarvioinnin kannalta oleellisia dimensioita voidaan luotettavasti mitata.

1.3 Psykologinen henkilöarviointi etäyhteydellä

Erilaiset digitaaliset ratkaisut, kuten etäyhteydellä tapahtuva psykologinen arviointi, tulevat mitä todennäköisimmin kasvattamaan merkitystään psykologian alalla nyt ja tulevaisuudessa. On tärkeää huomata, että psykologisella online- tai etäarvioinnilla voidaan viitata useampaan eri asiaan, kuten niin sanottuihin perinteisiin psykologisiin tutkimusmenetelmiin perustuvaan (esim. kognitiiviset kykytehtävät, inventaarit, haastattelu) asiantuntija-arvioon tai algoritmeille pohjautuvaan automatisoituun tietokonemenettelyyn. Tutkimuksissa on esimerkiksi havaittu, että digitaalisia jalanjälkiä, kuten Facebook-tykkäyksiä hyödyntämällä algoritmit voivat melko osuvasti arvioida niin henkilön persoonallisuutta kuin älykkyyttäkin (Kosinski ym. 2013; Youyou ym., 2015). Henkilöarvioinnin kontekstissa tulevaisuuden ratkaisuksi on esitetty myös erilaisia pelillistettyjä virtuaalisia simulaatioita, joissa henkilöt esiintyvät valitsemallaan hahmolla eli avatarilla (Howland ym., 2015). Tällaisessa virtuaalisessa ympäristössä henkilöistä voidaan kerätä suuri määrä arvioinnissa hyödynnettävää dataa aina tehtäväsuoriutumisesta hiiren liikautuksiin asti. On todennäköistä, että tekoäly tulee kasvattamaan rooliaan henkilöarviointityössä tulevaisuudessa ja että vähintäänkin menetelmät nojaavat entistä vahvemmin automatisoituun tietojenkäsittelyyn. Toistaiseksi tietokoneet eivät kuitenkaan ole syrjäyttäneet perinteistä henkilöarviointia, minkä vuoksi on tärkeää kiinnittää huomiota myös siihen, miten tavanomaiset arviointityökalut toimivat etäyhteydellä.

Vaikka yksittäisiä kysely- ja testimenetelmiä on tutkittu myös etäkontekstissa, on laajamittaisesta etähenkilöarvioinnista toistaiseksi vain vähän tutkittua tietoa. Tästä syystä keskityn tässä luvussa tarkastelemaan erilaisten arviointimenetelmien toimivuutta sekä psykologiseen arviointiin liittyviä erityispiirteitä silloin, kun prosessi toteutetaan kokonaisuudessaan etäyhteydellä. En kuitenkaan tutki tässä yhteydessä laajemmin esimerkiksi digitaalisiin jalanjälkiin ja tekoälyyn liittyviä sovellutuksia, vaan niin sanottuun perinteiseen henkilöarviointiin liittyviä tutkimusmenetelmiä etäteknologioita hyödyntäen. Tutkielmassa käytän etuliitettä sähköinen kuvaamaan menetelmiä, jotka yksilön on mahdollista suorittaa itsenäisesti esimerkiksi tietokoneella tai tabletilla.

1.3.1 Sähköisten arviointimenetelmien edut

Sähköisten psykologisten arviointimenetelmien käyttämiseen liittyy monia etuja, kuten hyvä saatavuus, joustavuus, helppous, automatisoitu tulosten tallentaminen, ajantasaisuus ja pienemmät kustannukset (Barak & Buchanan, 2004; Kaplan, 2018; Zygouris & Tsolaki, 2015).

Lisäksi useat tutkimukset ovat osoittaneet, että sähköiset arviointimenetelmät tuottavat yhtä luotettavia tai jopa luotettavampia tuloksia kuin perinteiset kynä-paperi-menetelmät (Barak & English, 2002; Campos ym., 2011; Meade ym., 2007; Rutherford ym., 2016; Sampson, 2000; Wall, 2000). Näin ollen sähköisten arviointimenetelmien suosiminen on monin tavoin perusteltua jopa ilman pandemiaa.

Kognitiiviset kykytestit kuuluvat käytetyimpien henkilöarviointimenetelmien joukkoon (Morris ym. 2015). Kognitiiviset kykytestit ovat tehtäviä, joilla kartoitetaan esimerkiksi kielellistä, numeerista ja avaruudellista päättelykykyä ja niiden on havaittu ennustavan työsuoritumista monissa yksittäisissä tutkimuksissa ja meta-analyyseissä (Arthur ym., 2003; Bobko ym., 1999; Hunter & Hunter, 1984; Morris ym., 2015; Schmidt ym., 2008; Schmidt & Hunter, 1998). Siihen, että arvioitava henkilö voi suorittaa kognitiiviset kykytehtävät sähköisesti etäyhteydellä, liittyy monia hyviä puolia (Kaplan, 2018). Yhtenä sähköisten kykytehtävien suurimpana hyötynä voidaan pitää sitä, että sähköinen esitysmuoto vähentää monia testaajasta johtuvia vinoumia testin teettämisessä ja inhimillisiä virheitä tehtävien pisteyttämisessä ja mitattavien muuttujien tallentamisessa. Lisäksi sähköiset menetelmät ovat kontrolloidumpia tilanteessa, jossa tehtävän tekijälle on annettava vain tietty määrä aikaa tai suoritukseen kuluva aikaa pitää mitata tarkasti. Kognitiivisten kykytehtävien onkin havaittu olevan toimivia myös etäkontekstissa (Geddes ym., 2020; Mooney, 2002; Rentz ym., 2016).

Myös inventaarit eli persoonallisuutta ja käyttäytymistä kartoittavat systemaattiset kyselymenetelmät ovat suosittuja menetelmiä henkilöarvioinnissa (Honkanen, 2005). Persoonallisuuden piirteiden arvioimista osana rekrytointia on aika ajoin kritisoitu, sillä yksittäisen persoonallisuuden piirteiden on meta-analyyseissä huomattu olevan melko heikosti yhteydessä työmenestykseen (Huang ym., 2014; Hurtz & Donovan, 2000; Morgeson ym., 2007; Morris ym., 2015; Ones & Viswesvaran, 1996; Schmidt ym., 2008). Toisaalta on havaittu, että viidestä suuresta piirteestä etenkin tunnollisuudella on voimaa työmenestyksen ennustajana, samoin kuin viittä suurta piirrettä tarkkarajaisemmilla ”alapiirteillä” sekä persoonallisuuden piirteiden yhdessä muodostamilla profiileilla (Barrick & Mount, 1991; Conte ym., 2017; Dudley ym., 2006; Huang ym., 2014; Tett ym., 2003). Tarkastelussa on syytä huomioda lisäksi se, että yhteydet persoonallisuuden ja työmenestyksen välillä voivat olla epälineaarisia, sillä esimerkiksi sekä hyvin matala että korkea tunnollisuus voi olla työelämän kannalta haitallista (Carter ym., 2014; Le ym., 2011). Myös henkilön kognitiivinen kyvykkyys voi vaikuttaa persoonallisuuden piirteen ilmenemiseen sillä henkilön kyky käyttää erilaisia ongelmanratkaisustrategioita esimerkiksi ahdistuksen hallintaan on riippuvainen kognitioista

(Perkins & Corr, 2006). Tärkeää on lopulta se, minkälaisesta kokonaisuudesta on kyse ja miten tietynlainen persoonallisuus ja työ sopivat yhteen (Mount, 1991; Dudley ym., 2006).

Se, että persoonallisuutta ja muita ominaisuuksia kartoittavia inventaareja sähköistetään, ei ole uusi ajatus. Etäyhteydellä toteutettujen inventaarien on havaittu olevan toimivia niin persoonallisuuden arvioimisessa (Cronk & West, 2002; Fouladi ym., 2002; Meade ym., 2007), luotettavuuskyselyissä (Jones ym., 2002) kuin uraan ja työhön liittyvissä mittareissa (Gati & Saka, 2001; Gore & Leuwerke, 2000; Oliver & Chartrand, 2000). On viitteitä jopa siitä, että ihmiset jakavat itsestään avoimemmin sensitiivistä informaatiota tietokonekyselyissä (Lin ym., 2016; Spear ym., 2016) ja että sosiaalista suotavuutta ilmenee tietokonekyselyissä vähemmän (Joinson, 1999). Hyvin korkea määrä sosiaalista suotavuutta voi olla ongelma muiden persoonallisuusdimensioiden luotettavuuden kannalta. Näin ollen sähköisten kyselymenetelmien käyttämiselle on monelta osin olemassa hyvät perusteet.

1.3.2 Sähköisiin arviointimenetelmiin liittyvät ongelmakohdat

Vaikka sähköisiin kysely- ja tehtävämenetelmiin liittyy kiistatta monia hyviä puolia, ei asia ole mutkaton. Tutkimustuloksia tarkasteltaessa on tärkeää huomioida ensinnäkin olosuhteiden merkitys ja menetelmää suorittavan henkilön motivaatio. Kliinisessä arvioinnissa henkilö on todennäköisemmin halukas antamaan itsestään totuudenmukaisen kuvan, mutta esimerkiksi rekrytointitilanteessa ihmiset pyrkivät usein luomaan itsestään todellisuutta paremman kuvan ja motivaatio huijaamiseen on ilmeinen (Bartram, 1997, 1999; Tippins ym., 2006). Erityisen suurena huolenaiheena huijaamista on pidetty kognitiivisten kykytehtävien kohdalla (Tippins ym., 2006). Huijaamisesta on kyse, jos esimerkiksi henkilöarviointiin osallistuva henkilö antaa kykytehtävät toisen tehtäväksi tai pyytää niihin apua. Sähköisten menetelmien suurena heikkoutena voidaankin pitää sitä, ettei henkilön identiteettiä ole välttämättä mahdollista todentaa. Huijaamisen kitkemiseksi on esitetty niin testituloksien varmentamista kontrolloiduissa oloissa kuin huijaamisen tunnistamiseen kehitettyjä data-analyysyjä (Tippins, 2009). On kuitenkin huomioitava, että vaikka huijaaminen on etäkontekstissa helpompaa, ei sitä välttämättä tapahdu kovinkaan suuressa mittakaavassa. Esimerkiksi Nye ja kumppanit (2008) eivät löytäneet merkkejä tämän kaltaisesta huijaamisesta omassa työnhakijoiden etäarviointia koskevassa tutkimuksessaan.

Toinen ongelma psykologisessa etäarvioinnissa on se, että suuri osa arviointimenetelmistä on suunniteltu kontrolloituihin oloihin, mutta etäkontekstissa tällaisia oloja ei voida taata (Barak & Buchanan, 2004). Ongelma on erityisen suuri sähköisten kykytehtävien kohdalla, sillä

esimerkiksi kotona syntyvää hälyä, henkilön vireystilaa tai keskeytyksiä ei voida kontrolloida tai havaita. Kaikki nämä ovat tekijöitä, jotka voivat vaikuttaa yksilön kykyyn keskittyä tai suoriutua tiedonkäsittelyä kuormittavissa tehtävissä. Lisäksi teknologia voi aiheuttaa muita yksilöstä riippumattomia ongelmia, kuten nettiyhteyden katkeaminen, mikä puolestaan asettaa tulosten hyödyntämisen kyseenalaiseksi. Tehtävien uudelleen tekeminen ei ole aina mielekäästä, sillä tuloksissa voi tällöin näkyä harjaantumisvaikutus.

Kolmas huoli sähköisissä arviointimenetelmissä liittyy siihen, että ihmiset ovat eritasoisia teknologioiden käyttäjiä ja taitamattomampien yksilöiden kohdalla tilanteesta syntyvä ahdistus voi vaikuttaa vastaustyyliin jopa persoonallisuusinventaaressa (Meade ym., 2007). Sähköisten arviointimenetelmien kohdalla on hyvin tärkeää, että näille menetelmille kerätään myös omat normit (Buchanan, 2003). Neljäs merkittävä huolenaihe on ollut testisalaisuuden menettäminen ja testi-itemien leviäminen hallitsemattomasti (Tippins ym., 2009). Lisäksi sähköisten arviointimenetelmien kohdalla huomioitavaa on myös se, että havainnot henkilön käyttäytymisestä jäävät usein tekemättä. Näin ollen esimerkiksi paineistumisesta, ohjeiden ymmärtämisestä, orientoitumisesta tai testitilanteeseen liittyvistä poikkeavuuksista ei voida kerätä informaatiota samaan tapaan kuin valvotun testitilanteen aikana, eikä tätä laadullista informaatiota voi taten hyödyntää tulosten tulkinnassa.

1.3.3 Vuorovaikutukselliset menetelmät psykologisessa etäarvioinnissa

Psykologisen etäarvioinnin ei tarvitse perustua pelkille sähköisille tehtäville ja kyselymenetelmille, sillä nykyiset teknologiat tarjoavat mahdollisuuden myös vuorovaikutukselle ja interaktiivisuudelle. Vaikka teknologiat ovat mahdollistaneet ihmisten välisen kanssakäymisen ja esimerkiksi psykologisten haastatteluiden ja henkilöarviointiin liittyvien simulaatioiden toteuttamisen myös etäyhteydellä, on tärkeää pohtia tämän vuorovaikutuksen laatua ja vaikutuksia.

Meta-analyysissään Blacksmith ja kumppanit (2016) havaitsivat ensinnäkin, että videoyhteyteen pohjautuvissa rekrytointihaastatteluissa haastattelijan haastateltavasta antama arvio on yleensä kielteisempi kuin kasvokkain tapahtuvassa haastattelussa (Blacksmith ym., 2016). Saman havaitsivat myös Basch ja kumppanit (2020) omassa tutkimuksessaan. Syitä tähän voivat olla haastateltavan heikompi mahdollisuus hyödyntää vaikutelman luomiseen tähtääviä strategioita (kuten visuaaliset vihjeet tai puhetapa), luoda katsekontaktia, ei-kielellisen informaation ja muiden sosiaalisten vihjeiden vähäisyys, vuorovaikutuksen rajoittuneisuus sekä viiveet äänessä ja kuvassa (Barrick ym., 2012; Blackman, 2002; Chapman

& Rowe, 2001; Ellis ym., 2002; Horn & Behrend, 2017; Stevens & Kristof 1995; Wegge, 2006). Aikaisemmissa tutkimuksissa on pantu merkille, että erilaiset vaikutelmanluontistrategiat ovat yhteydessä parempaan haastatteluarvioon (Barrick ym., 2009; Rosenfeld, 1997; Stevens & Kristof, 1995). Lisäksi on olemassa jonkin asteisia viitteitä siitä, että taipumus itsensä esittämiseen myönteisessä valossa on yhteydessä jopa parempaan esihenkilöarvioon työmenestyksestä (Ingold ym., 2015). Etäkommunikoinnin kannalta yksi mielenkiintoinen löydös on myös se, että niin sanottu picture in picture -tila, jossa haasteltava näkee haastattelijan lisäksi myös itsensä kuvaruudulla, lisää yleensä henkilön kognitiivista kuormittuneisuutta (Horn & Behrend, 2017). Tämä voi synnyttää negatiivisia tunteita, jotka puolestaan voivat vaikuttaa haastattelutilanteeseen ja haastatteluarvioon. Myös haastateltavat itse kokivat etähaastattelun negatiivisemmin kuin kasvokkain tapahtuvan haastattelun (Blacksmith ym., 2016).

Tutkimustulokset viittaavat siihen, että etäyhteyteen pohjautuva vuorovaikutus voi vaikuttaa yksilön käyttäytymiseen ja havaitsemisprosesseihin. Näin ollen etäyhteyden hyödyntämisellä voi olla vaikutuksia myös vuorovaikutteisiin henkilöarviointimenetelmiin kuten haastattelu sekä ryhmä- ja yksilösimulaatiot. Tutkimusten tuloksiin on kuitenkin suhtauduttava kriittisesti useammasta eri syystä. Ensinnäkin Blacksmithin ja kumppanien (2016) meta-analyysissä efektikoot eri tutkimuksissa olivat vaihtelevia. Mikäli tarkastellaan tilastollista merkitsevyyttä, kyseisen meta-analyysin tulos ei ylittänyt merkitsevyystasoa $p < .05$. Toiseksi on huomioitava se, että kommunikaatioteknologiat ovat kehittyneet ajan myötä yhä paremmin toimiviksi ja niiden käyttö on todennäköisesti tullut ihmisille yhä tutummaksi, mielekkäämmäksi sekä luontevammaksi. Kolmanneksi on mahdollista, että esimerkiksi monimenetelmällisyys ja henkilöarviointia tekevien henkilöiden asiantuntijuus paikkaavat vuorovaikutusalustasta syntyviä mahdollisia vinoumia ja puutteita henkilöarviointikokonaisuudessa.

On myös esitetty, että koska moni työnkuva sisältää tänä päivänä paljon etätyöskentelyä, etäyhteydellä toteutettu henkilöarviointi tavoittaa arvioitavat dimensiot jopa tarkoituksenmukaisemmin kuin perinteinen henkilöarviointi (Howland ym., 2015). Tämän tutkimuksen kannalta havainto on erityisen oleellinen, sillä pandemian vuoksi myös moni rekrytoinut yritys on toiminut pakotetusti kokonaan etäyhteyksien varassa. Selkeämpien vastausten saaminen etäyhteydellä toteutettavan psykologisen henkilöarvioinnin toimivuudesta edellyttää lisätutkimusta.

1.4 Tutkimusongelmat ja hypoteesit

Tässä tutkimuksessa verrattiin Psycon Oy:n henkilöarviointiaineistoa etähenkilöarviointitilanteesta ja perinteisen, lähikontaktissa toteutetun henkilöarvioinnin tilanteesta. Tutkimuksessa asetettiin kaksi hypoteesia.

Ensinnäkin haluttiin tutkia, kuinka hyvin etäyhteydellä toteutettu henkilöarviointi Psycon Oy:llä ennustaa todellista työsuoritumista esihenkilön arvioimana, verrattuna lähikontaktissa toteutettuun henkilöarviointiin. Psyconin henkilöarviointiprosessin perinpohjaisuuden, monimenetelmällisyyden, henkilöarviointiassistenttien ja arviointikonsulttien ammattitaidon sekä etäkontekstiin soveltuvien henkilöarviointimenetelmien vuoksi oletettiin, että etähenkilöarvioinneilla pystytään ennustamaan työmenestystä vähintään kohtalaisella tarkkuudella verrattuna perinteiseen henkilöarviointiin.

Toinen hypoteesi liittyi käytettyihin henkilöarviointimenetelmiin ja niiden avulla mitattaviin ominaisuuksiin. Tutkimuksessa selvitettiin, millä tavoin Psycon Oy:n käyttämät tehtävä- ja kyselymenetelmät sekä niiden mittaamat ominaisuudet ovat yhteydessä rekrytoidun henkilön työmenestykseen. Aiemman tutkimustiedon perusteella oletettiin, että esihenkilön arvioimaa työsuoritumista ennustavat ainakin kognitiivinen kyvykkyys ja jollain tasolla myös henkilön tunnollisuus. Tutkimustiedon perusteella ei ollut syytä olettaa, että suoritumista ennustavat ominaisuudet eroaisivat merkittävästi toisistaan eri tilanteiden välillä.

2 Menetelmät

2.1 Tutkimusasetelma ja tutkittavat

Psycon Oy:n etähenkilöarvioinnin seurantatutkimuksen aineisto kerättiin koronapandemian aikana toteutuneista, rekrytointien yhteydessä tehdyistä henkilöarvioinneista. Koronapandemiatilanteen muututtua vakavammaksi kaikki henkilöarvioinnit toteutettiin kokonaan etäyhteydellä maaliskuun 2020 puolivälin jälkeen. Jo ennen pandemiaa arviointisarja sisälsi tyypillisesti etäyhteydellä tehtäviä ennakkokyselyitä, mutta esimerkiksi simulaatiot, haastattelu ja kykytehtävät suoritettiin toimistolla. Ajankohta seurantatutkimuksessa rajattiin siten, että rekrytoitujen henkilöiden voitiin olettaa työskennelleen työtehtävässään yli 6 kuukautta siinä kohtaa, kun työmenestystä mitattiin, jotta menestymisen arvioiminen oli mielekästä.

Otanta koostui kaikista mukaan saaduista tarkastellulla aikavälillä Psyconin sopimusasiakkaille tehdyistä henkilöarvioinneista. Mukana oli laajasti eri alojen organisaatioita ja erilaisia työnkuvia. Perinteisen henkilöarvioinnin aineisto oli kerätty aikaisemmin vastaavanlaisena seurantatutkimuksena, jossa työntekijät olivat niin ikään työskennelleet työssään vähintään 6 kuukautta. Etähenkilöarviointiin osallistuneiden henkilöiden keski-ikä oli 39-vuotta, iän keskihajonnan ollessa 8.1. Tämä vastaa hyvin Psycon Oy:n henkilöarviointien tavanomaista ikärakennetta. Lähikontaktissa toteutuen henkilöarvioinnin osalta tarkka ikärakenne ei ollut tiedossa. Henkilön sukupuoli ei ollut muuttujana tässä tutkimuksessa, mutta Psyconin henkilöarviointeihin osallistuvista henkilöistä yleisesti noin 38 % on naisia ja 62 % miehiä.

Tutkimus toteutettiin Psycon Oy:n sisäisten ohjeiden, GDPR-asetuksen, tietosuojalainsäädännön ja asiakasorganisaatioiden kanssa luotujen yhteistyösopimusten mukaisesti. Tutkimuskäyttöä varten aineistosta poistettiin kaikki tunnistetiedot eli aineisto anonymisoitiin. Ennen anonymisointia henkilötietoja sisältävää aineistoa käsittelivät tietoturvasyistä vain seurantatutkimuksen toteuttamiseen nimetyt henkilöt Psycon Oy:n laitteilla. Tutkimukseen osallistuneiden organisaatioiden osalta lupa esihenkilöiden haastattelemiseen pyydettiin etukäteen organisaation HR-osastolta. Lisäksi kyselyyn vastaaminen oli jokaisen vastaajan osalta vapaaehtoista. Myös arvioitavilta henkilöiltä itseltään oli kysytty henkilöarviointiin osallistuessa lupaa heidän arviointiaineistonsa hyödyntämiseen osana Psycon Oy:n tutkimusta. Tässä tutkimuksessa ovat mukana vain luvan kirjallisessa muodossa antaneet arvioitavat. Organisaatiolle myös kerrottiin, että halutessaan he voivat

pyytää valmista anonyymiä dataa omaan tutkimuskäyttöön. Henkilötietoja sisältävän aineiston säilyttämisestä, luovuttamisesta ja tuhoamisesta vastaa aineiston rekisterinpitäjä eli Psycon Oy ohjeistuksien ja lainsäädännön edellyttämällä tavalla. Aineiston keräämisessä ja käsittelyssä oltiin erityisen huolellisia, jotta menettely vastaa sekä lainsäädäntöä, ohjeistuksia että myös Tutkimuseettisen neuvottelukunnan ihmiseen kohdistuvan tutkimuksen eettisiä periaatteita (Kohonen ym., 2019). Tutkimukseen ei pyydetty erikseen ihmistieteiden eettisen toimikunnan ennakkoarviota, sillä tutkimuksen toteutustapa ei sitä vaatinut.

2.2 Henkilöarviointimenetelmät

Psycon Oy:n etäarviointimalli on kehitetty jo ennen koronapandemiaa ja arvioinnissa yleisesti hyödynnetyt menetelmät on suunniteltu toimivaksi myös paikkariippumattomasti. Tyypillinen arviointisarja sisältää aina konsultin puolistrukturoidun haastattelun, kognitiivisen kyvykkyyden mittarit, persoonallisuus- ja toimintatyöli-inventaareja sekä ryhmäsimulaation. Näiden lisäksi sarjaan voidaan lisätä tarkemmin rajattuja kyselymenetelmiä, tarkkuus- ja organisointitehtäviä tai erilaisia yksilösimulaatioita. Myös Psycon Oy:n koulutetut assistentit tuottavat havaintoja arviointiin osallistuvien henkilöiden käyttäytymisestä. Eri menetelmien tuottaman informaation pohjalta psykologikonsultti tekee arvion, tuottaa raportin ja antaa myös numeerisen yleisarvion hakijan soveltuvuudesta työhön (asteikko 1-4, jossa 1= ”hän täyttää tehtävän asettamat vaatimukset ja valintaa voi suositella”, 2= ”hän täyttää tehtävän asettamat vaatimukset suurelta osin, joskin panostaminen kehittämisalueisiin parantaisi menestysennustetta”, 3= ”hän täyttää osan tehtävän asettamista vaatimuksista, ja kehittämisalueet on hyvä huomioida mahdollisina riskeinä” ja 4= ”hän ei vahvuuksistaan huolimatta täytä tehtävän asettamia vaatimuksia, joten valintaa ei voi suositella”).

Etätilanteessa kaikki tehtävät luonnollisesti suoritettiin etäyhteydellä. Myös läsnätilanteessa arvioitavat suorittivat tyypillisesti osan tehtävistä ennakkoon etäyhteydellä. Näihin tehtäviin kuuluivat esimerkiksi seuraavassa kappaleessa esiteltävä Motivaatiomittari, joka on toinen Psycon Oy:n tärkeimmistä persoonallisuus- ja toimintatyöli-inventaareista. Esimerkiksi kykytehtävät, simulaatiot ja konsultin haastattelu kuitenkin toteutettiin läsnätilanteessa Psycon Oy:n toimitiloissa. Tässä tutkimuksessa mahdollisten erojen persoonallisuuden piirteiden ennustavuudessa oletettiin liittyvän enemmän siihen, että etäaikakaudella tietyt piirteet korostuvat työmenestymisen kannalta, eikä siihen ympäristöön, jossa henkilö on vastannut kysymyksiin. Sekä kykytehtävien että inventaarien reliabiliteettiestimaattien laskemiseen on käytetty Cronbachin alfaa.

2.2.1 Kognitiiviset kykytehtävät

Kognitiivisen kyvykkyyden mittaaminen sisältää Psycon Oy:n henkilöarvioinnissa käytännöllisesti katsoen aina numeerisen päättelyn, kielellisen päättelyn ja visuaalis-loogisen päättelyn mittarit. Kaikki mittarit toimivat sähköisesti ja tehtäville on kerätty laajat normit useammalta tuhannelta Psycon Oy:n henkilöarviointiasiakkaalta ja niitä on validoitu eri tavoin (Furnham & Nederström, 2010; Psycon Oy, 2020). Normit on kerätty erikseen opisto- ja akateemisen tason hakijoille, mutta tutkimuksessa käytettiin kognitiivisten kykytehtävien raakapisteitä. Tässä osiossa esiteltyjen menetelmien reliabiliteettiestimaatit on kerätty Psyconin omasta aineistosta Ravenin progressiivisia matriiseja lukuun ottamatta (Raven ym., 1998).

Kolmesta kognitiivisesta mittarista Lukusarjat (Numeric Reasoning Test, Holzman ym., 1983) mittaa numeerista ja induktiivista päättelykykyä. Tehtävässä osallistujan tulee jatkaa lukusarjaa, kuten “1, 4, 7, 10, 13” kahdella luvulla sarjan logiikan mukaisesti. Psycon Oy:n oman version reliabiliteetti on .90 (Psycon Oy, 2020).

Kielellinen tiedon hyödyntäminen (Verbal Critical Reasoning Test, Klaczynski ym., 1997) mittaa kielellisloogista päättelykykyä. Menetelmä tarkastelee henkilön kykyä ajatella kriittisesti, rationaalisesti ja tehdä loogisia päätelmiä kielellisen tiedon avulla. Tehtävässä osallistujan tulee valita, onko esitetty väittämä tekstin perusteella tosi, epätosi vai eikö teksti sisällä riittävästi informaatiota päätelmän tekemiseen. Psycon Oy:n suomenkielisen version reliabiliteetti on .89 (Psycon Oy, 2017c).

Kolmas kognitiivisen kyvykkyyden mittari Psycon Ability Test (PAT) mittaa visuaalis-loogista päättelykykyä ja menetelmä on Psycon Oy:n kehittämä. Tehtävässä osallistujan tulee päätellä, millä tavalla kuviojono jatkuu. Tehtävän reliabiliteetti on .92 ja se on validoitu Psycon Oy:n muiden mittareiden ja esimerkiksi Ravenin progressiivisten matriisien suhteen (Psycon Oy, 2020; Raven ym., 1998). Lähikontaktissa toteutetun henkilöarvioinnin aineistossa PAT-testin tilalla käytettiin induktiivista päättelykykyä mittaavaa Ravenin progressiiviset matriisit -menetelmää (Raven’s Advanced Progressive Matrices, Raven ym., 1998). Tehtävässä osallistujan tulee päätellä, mikä kuvio sopii kuviomatriisin tyhjäksi jätettyyn kohtaan. Menetelmän reliabiliteettiä ovat välillä 0.83–0.87 (Raven ym., 1998). Psyconilla menetelmän normittaminen on perustunut yrityksen omaan henkilöarvioinnissa käyneiden asiakkaiden aineistoon, mutta menetelmän standardointiprosessi ja reliabiliteettiestimaattien hankkiminen on kuvattu tarkasti sen alkuperäisessä käsikirjassa (Psycon Oy, 2017c; Raven ym., 1998).

2.2.2 Motivaatiomittari

Motivaatiomittari (Nederström, 2008) on Psycon Oy:n kehittämä persoonallisuusinventaaari, jossa arvioitava henkilö arvioi erilaisten työ- ja arkikäyttäytymiseen liittyvien väittämien osuvuutta itseensä 4-portaisella Likert-asteikolla. Menetelmän avulla on tarkoitus kartoittaa erilaisia työkäyttöön vaikuttavia motivaatorakenteita ja sen viidentoista piirteen malli perustuu tutkittuun tietoon yleisimmistä työelämään ja -menestykseen liittyvistä piirteistä, tavoitteista ja motivaatiotekijöistä. Lisäksi mittari mittaa vastaustyylin sosiaalista suotavuutta. Taulukossa 1 on esitetty Motivaatiomittarin 17 piirrettä sekä niiden reliabiliteetit ja esimerkkipäätökset.

Taulukko 1. Motivaatiomittarin piirteiden nimet, reliabiliteetit ja esimerkkipäätökset

Motivaatiomittaripiirre	α	Esimerkkipäätös
Sosiaalinen tarve	.78	Aloitan usein keskusteluita spontaanisti vieraidenkin ihmisten kanssa
Luova tarve	.85	Olen parhaimmillani sellaisessa työssä, jossa toimitaan improvisoiden ja ilman selkeitä suunnitelmia
Johtamisen tarve	.83	Haluan välttää suoraa vallankäyttöä (käänteinen väittämä)
Jatkuvuuden tarve	.80	Arvostan säännöllistä ja ennakoitavaa elämää
Valmentamisen tarve	.76	Nautin saadessani ohjata ja opastaa tuttaviani
Tavoitteellisuuden tarve	.83	Olen valmis uhraamaan paljon, jotta ylittäisin muut ja itseni
Loppuunsaattamisen tarve	.79	Haluan saattaa vanhan projektin aina kunnolla loppuun, ennen kuin siirryn uusiin hankkeisiin
Huolestuvuus	.83	Jos epäonnistun työssäni, se vaivaa minua monta päivää
Empatia	.81	Joskus valvon öitä ystäväni ongelmien vuoksi
Riskinotto	.85	Teen joskus hyvinkin uhkarohkeita päätöksiä ongelmatilanteissa
Hyökkäävyys	.78	Valitan helposti ravintolassa, jos palvelussa on vikaa
Defensiivisyys	.83	Suhtaudun usein kriittisesti ihmisiin, kun tapaan heidät ensimmäistä kertaa
Optimismi	.75	Ennen pitkää ponnistelut ja oma aktiivisuus aina palkitaan
Tunnollisuus	.79	Siirrän hankalimmat tehtävät usein tulevaisuuteen (käänteinen väittämä)
Sisäinen kontrolliodotus	.81	Menestymisessä tärkeintä on, että on hyvää onnea ja sattuu olemaan oikeassa paikassa oikeaan aikaan (käänteinen väittämä)
Sosiaalinen suotavuus (itsepetos)	.80	En menetä ikinä malttiani ilman hyvää syytä
Sosiaalinen suotavuus (vaikutelmanluonti)	.80	Olen käyttänyt työpaikkani tarvikkeita ainoastaan työasioiden hoitamiseen

Jokaista Motivaatiomittarin piirrettä arvioidaan 9–12 väittämällä. Piirteiden reliabiliteettikeskiarvo on .80, yksittäisten piirteiden reliabiliteetin vaihdellessa .75 ja .85 välillä (Psycon Oy, 2017a). Menetelmä on validoitu Personality Research Formin (Niiamo, 1997), PK5-persoonallisuustestin (Tapaninen ym., 2007) sekä muiden Psycon Oy:n arviointimenetelmien suhteen. Lisäksi vastaustyylin sosiaalista suotavuutta kartoittavat kysymyksen on jaettu Paulhusin (1984) teorian mukaisesti kahteen aladimensioon: vähemmän tiedostettuun naiiviin ja rikkeettömään minäkuvaan (itsepetos) sekä paremmin tiedostettuun totuuden vääristelyyn (vaikutelmanluonti). Motivaatiomittari täytettiin etäyhteydellä molemmissa arviointikonteksteissa.

2.2.3 Stressireaktiokysely

Stressireaktiokysely (Stress Reaction Style Inventory, SRS; Nederström, 2007; Nederström & Furnham, 2012) on Psycon Oy:n kehittämä persoonallisuusinventaaari, joka mittaa kymmentä DSM-IV:n (Diagnostic and Statistical Manual of Mental Disorders) persoonallisuushäiriöihin perustuvaa piirrettä. Inventaarissa osallistuja arvioi erilaisten arki- ja työelämään liittyvien väittämien sopivuutta itseensä 4-portaisella Likert-asteikolla. Stressireaktiokyselyn nimi perustuu ajatukseen siitä, että persoonallisuuden patologisemmat piirteet nousevat esiin erityisesti stressitilanteissa. Mittari on kehitetty henkilöarviointitarkoitukseen, eikä sillä ole tarkoitus arvioida kliinisesti merkittävää persoonallisuuden patologiaa. Stressireaktiokyselyn piirteet, piirteiden reliabiliteetit ja esimerkkiväittämät on esitetty Taulukossa 2.

Taulukko 2. Stressireaktiokyselyn piirteiden nimet, reliabiliteetit ja esimerkkiväittämät

SRS-piirre	α	Esimerkkiväittämä
Eristäytyvyys	.76	Haluan pitää tiettyä etäisyyttä ystäväni ongelmiin
Omintakeisuus	.79	Huumorintajuani pidetään usein erikoisena
Epäsosiaalisuus	.68	Tuloksellinen toiminta oikeuttaa kovatkin otteet
Äkkijyrkkyys	.76	Räjähdän usein pikkuasioiden takia, mitä myöhemmin kadun
Huomionhakuisuus	.87	Pidän siitä, että saan ihmisten huomion käännettyä itseeni
Narsismi	.74	Olen monissa tilanteissa itseoikeutettu johtaja
Estyneisyys	.71	Suuressa joukossa mietin tarkkaan, mitä uskallan sanoa ääneen
Riippuvaisuus	.70	Ryhmässä mielelläni päätän, miten asioissa kannattaisi edetä (käänteinen väittämä)
Pedanttisuus	.76	Haluan asettaa työlleni mahdollisimman tarkan suunnitelman ja aikataulun
Sosiaalinen suotavuus	.84	En koskaan ajattele pahaa työtovereistani

SRS = Stress Reaction Style

Stressireaktiokyselyn jokaista kymmentä piirrettä kartoitetaan neljällätoista itsenäisellä väittämällä. Kokonaisuudessaan kysely sisältää 152 väittämää, joiden reliabiliteetti vaihtelee välillä .68 ja .87 (Psycon Oy, 2017b). Lisäksi kysely mittaa vastaajan sosiaalista suotavuutta 12 väittämällä. Menetelmä on ristiin validoitu Personality Research Formin kanssa (Niitamo, 1997). Stressireaktiokysely on täytetty tyypillisesti toimistolla lähikontaktissa toteutetun henkilöarvioinnin osalta.

2.3 Toteutunut työmenestys

Etähenkilöarvioinnin yhteydessä toteutunutta työmenestystä tutkittiin lyhyellä kyselyllä. Osa vastauksista kerättiin puhelinhaastatteluilla ja osa sähköisillä kyselylomakkeilla. Ensisijainen tiedonkeruutapa oli puhelinhaastattelu, mutta organisaation tai henkilön toiveesta esihenkilöiden oli mahdollista vastata myös sähköiseen lomakkeeseen. Tiedonkeruussa suosittiin puhelinhaastattelua, jotta katoa ilmenisi vähemmän. Kyselyiden sisältö oli identtinen, eikä kyselyn muodon oletettu vaikuttavan vastaustapaan.

Kyselyssä kuhunkin rekrytointiin liitettyltä lähiesihenkilöltä kysyttiin, kuka henkilöarvioinnin läpikäyneistä henkilöistä oli valittu työhön ja kuinka hyvin henkilö on pärjännyt työssään. Esihenkilön arvioimaa työsuoriutumista selvitettiin viidellä kysymyksellä, joihin vastaus annettiin 4-portaisella Likert-asteikolla. Sama asteikko oli käytössä myös lähikontaktissa toteutetun henkilöarvioinnin seurantatutkimuksessa. Kysymyksillä kartoitettiin esimerkiksi työmenestystä, työssä viihtymistä ja sosiaalisten taitojen vastaavuutta työn vaatimuksiin. Tämän lisäksi esihenkilöltä kysyttiin kaksi avointa kysymystä, joissa kartoitettiin työsuoriutumiseen vaikuttaneita, henkilöstä riippumattomia tekijöitä sekä pyydettiin kommentteja arviointiprosessista sekä Psycon Oy:stä yleensä. Näin ollen vastaajien oli mahdollista tehdä täsmennyksiä vastauksiinsa myös verbaalisesti. Tässä tutkimuksessa tarkasteluun otettiin numeerinen aineisto ja työmenestyksen vastemuuttujana käytettiin kysymystä ”kuinka hyvin henkilö on menestynyt työssään”. Lisäksi kerättiin kontrollimuuttujana tietoa siitä, kuinka luotettavaksi esihenkilö arvioi omat vastauksensa. Kyselyjen tarkka sisältö ja muotoilu ovat Psycon Oy:n omaisuutta, eikä niitä tästä syystä ole sellaisenaan raportoitu tämän tutkimuksen liitteissä. Läsnä toteutetun henkilöarvioinnin aineisto sisälsi samat seurantakysymykset ja myös lähes täysin samat arviointimenetelmät, minkä johdosta aineistojen vertailu oli näiltä osin mahdollista.

2.4 Analyysit ja kriteerit

Henkilöt, joilta puuttui henkilöarvioinnin yleisarvio, poistettiin aineistosta, sillä näiden henkilöiden kohdalla ei olisi ollut mahdollista arvioida, onko työmenestystä onnistuttu ennustamaan. Tällä periaatteella tuoreesta etäarviointiaineistosta poistettiin 5 osallistujaa. Lisäksi yksi henkilö oli irtisanoutunut itse koeajalla ja toinen vaihtoi koeajalla työpaikkaa, minkä vuoksi nämä henkilöt jäivät tämän tutkimuksen ulkopuolelle. Lähikontaktissa toteutetun henkilöarvioinnin aineistossa vähintään 6 kk työssään olleita osallistujia, joiden osalta oli saatavilla tieto sekä työmenestyksestä että henkilöarvioinnin suosituksesta, oli 149:stä yhteensä 91. Aineistosta jätettiin pois myös yksi henkilö, jonka työmenestyksen arviointia ei voitu pitää luotettavana hänen henkilökohtaiseen elämäänsä liittyvien syiden vuoksi. Vanhan aineiston osalta suurin syy puuttuville havainnoille on rajattu aika henkilöarvioinnin tulosten säilyttämiselle. Toinen syy liittyy siihen, että korkeimman johtotason tehtävissä ei välttämättä anneta numeerista suositusta. Koska kato oli suurelta osin satunnaista, ei sen katsottu vaikuttavan tämän tutkimuksen keskeisiin tuloksiin. Myöskään silmämääräinen tarkastelu ei antanut aihetta huoleen tutkimustulosten vinoutumisen suhteen.

Tilastolliset analyysit suoritettiin IBM SPSS Statistics 27 -ohjelmalla (IBM Statistical Package for the Social Sciences). Aluksi henkilöarvioinnin yleisarvion skaala käännettiin, jotta se oli saman suuntainen esihenkilöiden antamien arvioiden kanssa. Tällöin heikoin arvio oli 1 ja paras arvio 4. Analyysijä jatkettiin tarkastelemalla muuttujien kuvailevia tunnuslukuja, jakaumia, frekvenssejä sekä muuttujien välisiä korrelaatioita.

Havainnot henkilöiden työmenestyksestä eivät olleet jakautuneet normaalisti ja vaihteluvälin kapeutuminen näkyi aineistossa. Tämän vuoksi lineaarisuuteen perustuvat tilastomenetelmät eivät tulleet kyseeseen työmenestyksen analysoinnissa. Henkilöarviointien osuvuutta arvioitiin henkilöarvioinnin suosituksen ja esihenkilön arvioiman työmenestyksen erotuspistemäärien jakaumien avulla. Samaa tarkastelutapaa käytettiin myös Heikkilän (2017) tutkimuksessa ja aikaisemmissa Psycon Oy:n sisäisissä seurantatutkimuksissa. Jakaumien eroa testattiin khiin neliö -testillä. Lisäksi tarkasteltiin, eroaako esihenkilöarvioiden luotettavuus ja arviointiraporttien kantaottavuus eri tilanteissa toisistaan niin ikään khiin neliö -testillä. Lopuksi katsottiin, vaikuttaako haetun tehtävän taso (perus/asiantuntija/esimies- ja päällikkö/johtaja) arvioinnin osuvuuteen etätilanteessa, jossa kyseinen tieto oli käytettävissä. Jakaumien eroa testattiin khiin neliö -testillä.

Tämän jälkeen tutkittiin, miten teorian pohjalta työmenestykseen yhteydessä olevat muuttujat eli kognitiivinen kyvykkyys ja tunnollisuus ovat yhteydessä esihenkilön arvioimaan työmenestykseen. Tarkasteluun otettiin mukaan ne kognitiiviset kykytehtävät, jotka olivat osa arviointisarjaa sekä etä- että läsnätilanteessa. Nämä tehtävät olivat numeerisinduktiivisen päättelyn tehtävä ”Lukusarjat” (Numeric Reasoning Test, Holzman ym., 1983) sekä kielellisen päättelyn tehtävä ”Kielellinen tiedon hyödyntäminen” (Verbal Critical Reasoning Test, Klaczynski ym., 1997). Asiaa tutkittiin muodostamalla logistinen regressiomalli. Logistiseen regressioon päädyttiin siitä syystä, että peräti 72 % rekrytoiduista henkilöistä oli saanut esihenkilöltään parhaan arvion työmenestyksestä ja näin ollen työmenestystä oli perustellumpaa tarkastella dikotomisena kuin jatkuvana muuttujana tässä aineistossa. Parhaan esihenkilöarvion saaneet henkilöt saivat arvon 1, joka tarkoitti ”on menestynyt työssään erinomaisesti” ja muut arvon 0, joka tarkoitti ”ei ole menestynyt työssään erinomaisesti”.

Koska kyseessä oli vertaileva asetelma, tutkittiin myös sitä, eroaako menetelmien kyky ennustaa työmenestystä etä- ja läsnätilanteessa. Tätä tutkittiin siten, että kustakin tarkastelun kohteena olevasta muuttujasta muodostettiin oma logistinen regressiomalli, joka sisälsi kyseisen muuttujan, tilanteen (etä/läsnä), joka koodattiin kategoriseksi muuttujaksi sekä muuttujan ja tilanteen välisen yhdysvaikutuksen. Mikäli yhdysmuuttujan tilastollinen merkitsevyys oli $p < .10$, voitiin päätellä, että tilanteella on vaikutusta muuttujan kykyyn ennustaa työmenestystä. Mikäli tämä ehto ei kuitenkaan täyttynyt, voitiin olettaa, että muuttuja toimii yhtä hyvin molemmissa tilanteissa ja koko aineiston voi kyseisen muuttujan osalta analysoida yhdessä. Päävaikutusten tarkastelussa raja-arvona oli $p < .05$. Tarkastelut suoritettiin erikseen kaikille tarkastelun kohteena oleville persoonallisuuden piirteille sekä kognitiivisen kyvykkyuden tehtäville. Lopuksi muodostettiin logistinen regressiomalli, jossa mukana olivat kaikki teorian pohjalta valitut muuttujat ja vasteena esihenkilön dikotomisoitu arvio työmenestyksestä.

Teorian pohjalta laaditun regressiomallin lisäksi muodostettiin aineistolähtöinen logistinen regressiomalli, jossa mukana tarkastelussa olivat kaikki persoonallisuuden piirteet ja kognitiiviset tehtävät. Näin ollen tarkasteltuja menetelmämuuttujia oli mukana yhteensä 32. Kunkin ominaisuuden kykyä ennustaa työsuoritumista tarkasteltiin yksittäin ja myös mahdollinen arviointikontekstin vaikutus ominaisuuden ennustavuuteen huomioitiin edellä kuvatulla tavalla. Lopuksi muodostettiin parhaiten ennustavista arviointimenetelmien dimensioista oma logistinen regressiomalli etäarvioinnille ja läsnäarvioinnille. Malleista karsittiin pois ne muuttujat, jotka eivät olleet tilastollisesti merkitseviä.

3 Tulokset

3.1 Henkilöarvioinnin yleisarvion ennustavuus

Henkilöarvioinnin yleisarvion ja esihenkilön arvioiman työmenestyksen kuvailevat tunnusluvut on esitetty Taulukossa 3. Kummassakaan arviointikontekstissa yhtäkään henkilöä ei rekrytoitu heikoimmalla henkilöarvioinnin suosituksella ja peräti 94 % kaikista oli rekrytoitu joko parhaalla (4) tai toiseksi parhaalla (3) henkilöarvioinnin suosituksella. Vaihtelua oli vähän myös esihenkilöiden tekemissä arvioissa. Kaiken kaikkiaan 96 % rekrytoiduista henkilöistä sai joko parhaan tai toiseksi parhaan arvion omalta esihenkilöltään. Henkilöarvioinnin yleisarvioiden ja esihenkilöarvioiden frekvenssit on raportoitu Taulukossa 4.

Taulukko 3. Henkilöarvioinnin yleisarvion ja esihenkilöarvioiden keskiarvot, keskihajonnat ja otoskoot etä- ja läsnätilanteessa

Muuttujat	Etäyhteys			Läsnä		
	ka	kh	n	ka	kh	n
Käännetty yleisarvio	3.52	.57	96	3.40	.61	90
Esihenkilöarvio	3.67	.56	96	3.60	.68	90

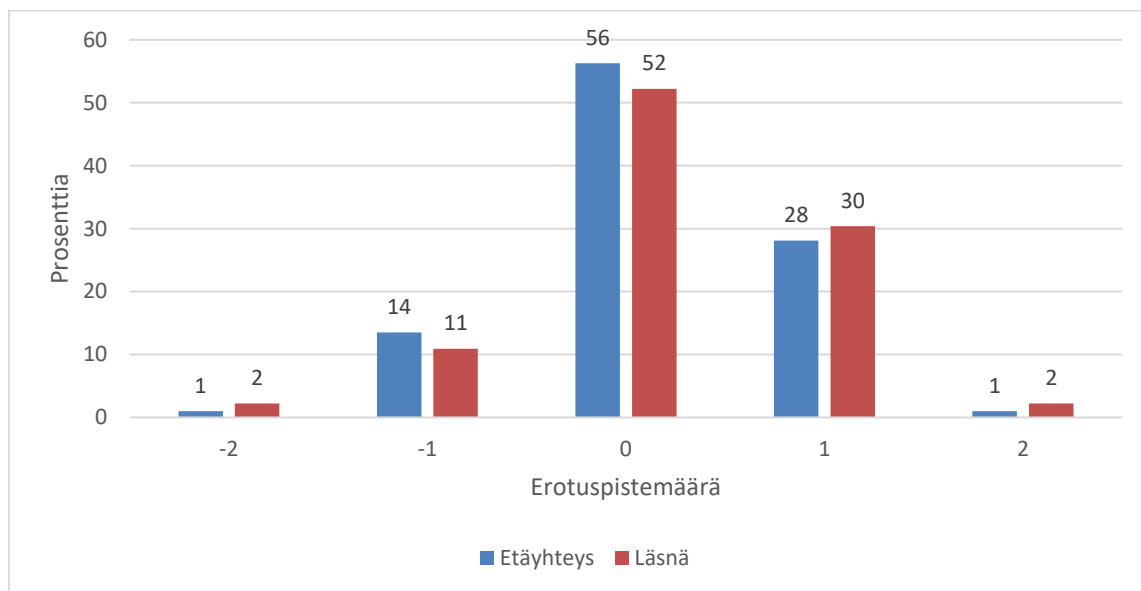
ka = keskiarvo, kh = keskihajonta, n= otoskoko

Taulukko 4. Esihenkilön arvioiman työmenestyksen ja henkilöarvioinnin yleisarvioiden ja frekvenssit ristiintaulukoituna (1 = heikoin arvio – 4 = paras arvio)

	Esihenkilöarvio	Henkilöarvioinnin yleisarvio				
		1	2	3	4	Kaikki
Etäyhteys	1	0	1	0	0	1
	2	0	1	0	1	2
	3	0	0	13	12	25
	4	0	1	27	40	68
	Kaikki	0	3	40	53	96
Läsnä	1	0	1	2	0	3
	2	0	0	1	0	1
	3	0	3	14	8	25
	4	0	2	25	34	61
	Kaikki	0	6	42	42	90

Henkilöarviointien osuvuutta tarkasteltiin esihenkilöiden suoritusarvioiden ja arviointisuositusten erotuspistemäärien avulla. Sekä etätilanteessa että läsnätilanteessa yli 50 % Psycon Oy:n suosituksista ja esihenkilön arvioista vastasi toisiaan täysin, eli suosituksen ja esihenkilöarvion erotus oli 0. Tilanteita, joissa suositus oli ollut yhden pisteen esihenkilön

antamaa arviota kriittisempi, oli etätilanteessa 28 % ja läsnätilanteessa 30 %. Tilanteita, joissa esihenkilön arvio oli puolestaan yhden pisteen verran henkilöarvioinnin suositusta kriittisempi, oli etätilanteessa 14 % ja läsnätilanteessa 11 %. Kahden pisteen eroja oli molemmissa aineistossa hyvin vähän. Näin ollen etätilanteessa arviot vastasivat toisiaan täysin tai lähes täysin 98 % ja läsnätilanteessa 94 %. Nämä jakaumat on raportoitu Kuvassa 1. Erotuspistemäärien tilastollista eroa etätilanteen ja läsnätilanteen välillä tarkasteltiin vielä khiin neliö -testillä ja havaittiin, että tilanteella ei ollut yhteyttä arviointiraportin suosituksen osuvuuteen $\chi^2(4) = 1.24$, $p = .87$. Kun tarkasteltiin erikseen, onko etä- ja läsnätilanteen välillä eroa siinä, kuinka paljon esiintyy selkeää epäsuhtaa (erotuspistemäärä yli 1 piste) arvioiden välillä, eroa ei havaittu $\chi^2(1) = .83$, $p = .36$.



Kuva 1. Työmenestyksen ja käännetyn henkilöarvioinnin yleisarvion erotuspistemäärien prosentuaaliset osuudet etätilanteessa ja läsnätilanteessa

Esihenkilöiden vastaukset oman arvionsa luotettavuudesta eivät eronneet etä- ja läsnätilanteessa tilastollisesti merkitsevästi toisistaan $\chi^2 = 3.97$, $p = .14$. Etätilanteessa 96 % ja läsnätilanteessa 100 % vastaajista arvioi oman vastauksensa luotettavaksi tai erittäin luotettavaksi. Erittäin luotettavissa (vastaus 4) vastauksissa menestyminen arvioitiin keskimäärin paremmaksi kuin luotettavissa vastauksissa (vastaus 3) $\chi^2(6) = 16.23$, $p = .01$. Tilastollista eroa konsulttien antamien suositusten määrässä ei ollut etätilanteessa ja läsnätilanteessa, vaan raportit ottivat kantaa yhtä paljon molemmissa konteksteissa $\chi^2(2) = 2.13$, $p = .35$. Toisin sanoen konsultit antoivat yhtä rohkeasti parhaan suosituksen molemmissa tilanteissa. Arvioinnin taso etätilanteessa ei vaikuttanut sen osuvuuteen $\chi^2(12) = 7.7$, $p = .81$.

Aineistossa perustason arviointeja oli 4 %, asiantuntija-arviointeja 51 %, esimies- ja päällikköarviointeja 42 % ja johtaja-arviointeja 3 %.

3.2 Teorialähtöinen logistinen regressioanalyysi

Teorian pohjalta oletettiin, että erinomaiseen työmenestykseen ovat yhteydessä henkilön kognitiivisen kyvykkyyden osa-alueet sekä tunnollisuus. Teorialähtöisen regressioanalyysin regressiokertoimet, niiden keskihajonnat, 95 % luottamusvälit efektikoolle sekä standardoidut selitysasteet eri tilanteissa on raportoitu Taulukossa 5. Taulukossa esitetyt kertoimet on laskettu siten, että kunkin muuttujan yhteyttä työmenestykseen on tarkasteltu yksittäin.

Taulukko 5. Teorialähtöisen logistisen regressioanalyysin keskeiset tulokset eri tilanteissa

Tilanne	Muuttuja	B	SE	OR (CI 95%)	R2
Etäyhteys	Tunnollisuus	-0.19	0.11	0.83 (0.67–1.03)	0.05
	Lukusarjat	0.11**	0.04	1.12 (1.03–1.22)	0.11
	KTH	0.03	0.03	1.03 (0.97–1.08)	0.01
	Psycon Ability Test	0.12	0.04	1.01 (0.94–1.08)	0.00
Läsnä	Tunnollisuus	0.15	0.11	1.17 (0.93–1.46)	0.03
	Lukusarjat	0.07	0.04	1.07 (1.00–1.16)	0.06
	KTH	0.04	0.03	1.04 (0.98–1.11)	0.03
	Raven	0.12*	0.60	1.13 (1.01–1.27)	0.07
Koko aineisto	Tunnollisuus	-0.03	0.08	0.97 (0.84–1.13)	0.00
	Lukusarjat	0.90**	0.03	1.10 (1.04–1.16)	0.08
	KTH	0.03	0.02	1.03 (1.00–1.08)	0.02

* $p \leq .05$ ** $p \leq .01$ *** $p \leq .001$

KTH = Kielellinen tiedon hyödyntäminen

B = Regressiokerroin

R2 = Nagelkerke R2

SE = Keskivirhe

OR = Odds ratio

Teorian pohjalta tarkasteluun valituista muuttujista numeerisinduktiivinen päättely ennusti työmenestystä koko aineistossa ($B = .90$, $p = .03$). Tilanteen ja menetelmän välistä yhdysvaikutusta ei havaittu. Tunnollisuuden osalta löydettiin tilanteen ja piirteen välinen yhdysvaikutus, jossa tunnollisuus näytti ennustavan työmenestystä negatiivisesti etätilanteessa mutta positiivisesti läsnätilanteessa ($B = -.34$, $p = .03$). Päävaikutuksia ei kuitenkaan havaittu, eli tunnollisuus ei lopulta ennustanut työmenestystä kummassakaan tilanteessa (etätilanteessa $B = -.19$, $p = .09$; läsnätilanteessa $B = .15$, $p = .19$). Kielellisen päättelyn osalta ei löydetty pää-

eikä yhdysvaikutuksia. Kielellinen päättely ennusti työmenestystä heikosti tässä aineistossa ($B = .03, p = .10$)

Yksittäisten tarkastelujen lisäksi tutkittiin, kuinka hyvin useamman muuttujat regressiomalli ennustaa erinomaista työmenestystä. Teorian pohjalta rakennettu logistinen regressiomalli, jossa oli mukana tunnollisuus, numeerisinduktiivinen päättely ja kielellinen päättely, ennusti työmenestystä, $\chi^2(3) = 11.74, p = .01$. Malli luokitteli oikein 67,2 % tapauksista ja sen selitysaste oli 9 % (Nagelkerke $R^2 = .09$). Numeerisinduktiivisessa päättelyssä menestyminen lisäsi todennäköisyyttä kuulua erinomaisesti menestyneisiin 1.1-kertaiseksi [1.04, 1.16]. Muut mittarit eivät ennustaneet mallissa erinomaista työmenestystä.

3.3 Aineistolähtöinen logistinen regressioanalyysi

Aineistolähtöisessä logistisessa regressioanalyysissä erinomaista työmenestystä ennustivat numeerisinduktiivinen päättely ($B = .90, p = .03$) ja läsnätilanteessa käytetty menetelmä Ravenin progressiiviset matriisit (Raven ym., 1998), joka mittaa induktiivista päättelykykyä ($B = .12, p = .04$). Sen sijaan henkilön defensiivisyys ($B = -.23, p = .02$) ja estyneisyys ($B = -.08, p = .01$) olivat käänteisessä yhteydessä erinomaiseen työmenestykseen koko aineistossa. Tilanteen ja psykologisten ominaisuuksien välisiä yhdysvaikutuksia ei havaittu, eli ominaisuuden kyky ennustaa työmenestystä ei riippunut tilanteesta. Kaiken kaikkiaan tarkasteltiin 28 persoonallisuuden ulottuvuutta ja 4 kognitiivisen kyvykkyyden osa-aluetta, mutta suurin osa dimensioista ei ollut yhteydessä työmenestykseen. Yksittäin tarkasteltuna aineistosta esiin nousseiden muuttujien regressiokertoimet ja muut kuvailevat tiedot on raportoitu Taulukossa 6.

Parhaiden työmenestyksen ennustajien pohjalta rakennettiin oma useamman muuttujan logistinen regressiomalli etätilanteelle ja läsnätilanteelle. Etätilanteessa malli, jossa kaikki muuttujat olivat yhdessä tarkasteltuna tilastollisesti merkitseviä, sisälsi numeerisinduktiivisen päättelyn ja estyneisyyden. Malli ennusti erinomaista työmenestystä, $\chi^2(2) = 14.97, p < .001$. Malli luokitteli oikein 75.5 % tapauksista ja sen selitysaste oli 21 % (Nagelkerke $R^2 = .21$). Numeerisinduktiivisessa päättelyssä menestyminen lisäsi todennäköisyyttä kuulua erinomaisesti menestyneisiin 1.13-kertaiseksi [1.04, 1.23] ja estyneisyys heikensi todennäköisyyttä kuulua erinomaisesti menestyneiden ryhmään 0.88-kertaiseksi [0.82, 0.97]. Läsnätilanteessa regressiomalliin jäi tilastollisen karsinnan jälkeen ainoastaan Ravenin progressiiviset matriisit (Raven ym., 1998). Malli ennusti erinomaista työmenestystä, $\chi^2(1) = 4.55, p = .03$. Malli, jossa oli pelkästään Raven luokitteli oikein 69.8 % tapauksista ja sen

selityssaste oli 7 % (Nagelkerke $R^2 = .07$). Ravenissa menestyminen lisäsi todennäköisyyttä kuulua erinomaisesti menestyneisiin 1.13-kertaiseksi [1.00, 1.27]. Multikollineaarisuusongelmaa ei ollut tarkasteltujen muuttujien suhteen.

Taulukko 6. Aineistolähtöisen logistisen regressioanalyysin keskeiset tulokset eri tilanteissa

Tilanne	Muuttuja	B	SE	OR (CI 95%)	R ²
Etäyhteys	Estyneyisyys	-0.11*	0.04	0.90 (0.82–0.98)	0.10
	Defensiivisyys	-0.21	0.12	0.82 (0.64–1.03)	0.04
	Lukusarjat	0.11**	0.04	1.12 (1.03–1.22)	0.11
Läsnä	Estyneyisyys	-0.05	0.05	0.95 (0.86–1.05)	0.02
	Defensiivisyys	-0.27	0.16	0.76 (0.56–1.04)	0.05
	Lukusarjat	0.07	0.04	1.07 (1.00–1.16)	0.06
	Raven	0.12*	0.60	1.13 (1.01–1.27)	0.07
Koko aineisto	Estyneyisyys	-0.08**	0.03	0.92 (0.86–0.98)	0.05
	Defensiivisyys	-0.23*	.10	0.80 (0.66–0.96)	0.04
	Lukusarjat	0.90**	0.03	1.10 (1.04–1.16)	0.08

* $p \leq .05$ ** $p \leq .01$ *** $p \leq .001$

KTH = Kielellinen tiedon hyödyntäminen

B = Regressiokerroin

R² = Nagelkerke R²

SE = Keskivirhe

OR = Odds ratio

4 Pohdinta

Tässä tutkimuksessa kokonaan etäyhteydellä toteutettu henkilöarviointi ennusti esihenkilön arvioimaa rekrytoidun henkilön työmenestystä yhtä hyvin kuin lähikontaktissa toteutettu henkilöarviointi Psycon Oy:llä. Sekä etäyhteydellä että lähikontaktissa toteutetuissa henkilöarvioinneissa yleisarvio ja esihenkilöarvio vastasivat hyvin toisiaan; 96 % tapauksista erotuspistemäärä esihenkilöarvion ja henkilöarvioinnin yleisarvion välillä oli maksimissaan 1 piste. Yksittäisten menetelmien ennustavuus jäi heikoksi, mikä tukee ajatusta siitä, että yksittäinen ominaisuus on harvoin universaalisti hyvä tai huono työmenestyksen kannalta. Yksittäisistä menetelmistä numeerinen ja induktiivinen päättelykyky ennustivat parempaa työmenestystä, kun henkilön defensiivisyys ja estyneisyys puolestaan heikensivät menetyssennustetta. Tilanne (etäyhteys/läsnä) ei vaikuttanut olennaisesti menetelmien kykyyn ennustaa työmenestystä.

4.1 Henkilöarvioinnin osuvuus

Tämä tutkimus on tietävästi ensimmäinen kokonaan etäyhteydellä toteutetun henkilöarvioinnin seurantatutkimus, jossa käytetyt menetelmät eivät liity pelillistämiseen tai tekoälyyn, minkä vuoksi tutkimuksen tuloksia ei voi suoraan verrata aiempaan tutkimustietoon. Koska etäyhteydellä toteutetussa psykologisessa arvioinnissa hyödynnettävien sähköisten arviointimenetelmien käyttöön liittyy tutkimuskirjallisuuden perusteella kuitenkin selviä riskejä, voidaan näin voimakasta osuvuutta etähenkilöarvioinneissa pitää pienoisenä yllätyksenä (Barak & Buchanan, 2004; Barrick ym., 2009; Blacksmith ym., 2016; Meade ym., 2007; Tippins ym., 2006). Etäyhteyden käyttöön liittyen aikaisemmin huolenaiheiksi nousseet teemat, kuten huijaaminen, erot vuorovaikutuksen laadussa, kriittisemmät haastatteluarviot, havaintojen vähäisyys, kontrolloitujen olosuhteiden puute tai mahdollinen häly ja muut häiriötekijät eivät ainakaan systemaattisesti vaikuttaneet henkilöarviointiraporttien osuvuuteen.

Todennäköisesti yksi tärkeä osa onnistunutta etähenkilöarviointia on monimenetelmällisyys, jonka on läsnä toteutetussa henkilöarvioinnissa havaittu parantavan arvioinnin laatua (Woehr & Arthur, 2003). Psycon Oy:n assistentit ovat tuottaneet arvioitavista henkilöistä laadullista informaatiota myös etäkontekstissa, arviointi on sisältänyt erilaisia soveltavia tehtäviä, eikä henkilöarviointi näin ollen ole perustunut pelkille menetelmien antamille tunnusluvuille ja konsultin haastattelulle myöskään etätilanteessa. Erilaisia simulaatiotehtäviä on myös ollut mahdollista hyödyntää aikaisempaa enemmän ajankäytön tehostumisen myötä. Näin ollen

laadullista informaatiota ja havaintoja on ollut mahdollista kerätä myös etähenkilöarvioinnin yhteydessä. Yhtenä monimenetelmällisyyden muotona voidaankin pitää myös sitä, että laadullista informaatiota tuottavat useammat eri henkilöt, jolloin yksilön subjektiiviset vinoumat eivät vaikuta kokonaistulkintaan yhtä voimakkaasti kuin tilanteessa, jossa havaintoja tuottaa vain yksi henkilö.

Koska etätyö ja digitaaliset toimintaympäristöt ovat olennainen osa nykyaikaista työelämää, voidaan nyt saatuja tuloksia etähenkilöarvioinnin toimivuudesta pitää tärkeinä. Tämän tutkimuksen perusteella etähenkilöarviointi vaikuttaa olevan hyvin toteutettuna vähintään yhtä hyvin toimiva malli työmenestyksen ennustavuuden suhteen kuin perinteinen henkilöarviointi, ainakin silloin, kun myös suoritettava työ tapahtuu voittopuolisesti etäyhteydellä. Kaiken kaikkiaan henkilöarviointien hyvä osuvuus tukee ajatusta menettelyn hyödyllisyydestä osana rekrytointia. Lisäksi psykologinen arviointi on keskeinen osa monen psykologin työnkuvaa ja uusi tutkimustieto etäarvioinnin toimivuudesta voi rohkaista koko alaa uudistumaan ja muovautumaan 2020-luvun tarpeisiin.

Johtopäätösten tekemisessä on oltava mukana myös kriittisyyttä muun muassa seuraavaksi esitellyistä syistä ja myöhemmin käsiteltävien tutkimukseen liittyvien rajoitusten vuoksi. Ensinnäkin on mahdollista, että tutkimustuloksiin ovat vaikuttaneet erilaiset väliin tulevat tekijät, sillä henkilön työmenestys ja sen arvioiminen ovat monen keskenään vuorovaikutuksessa olevan asian summa. Väliin tulevia tekijöitä voivat olla esimerkiksi vahvistusharha, joka viittaa yksilön taipumukseen kiinnittää enemmän huomioita omaa ennakkokäsitystä tukevaan informaatioon (Jones, 2001). Näin ollen esihenkilö saattaa kiinnittää enemmän huomioita hyvän suosituksen saaneen työntekijän positiivisiin ominaisuuksiin kuin kehitysalueisiin. Lisäksi on huomioitava, että jotkut piirteet henkilössä saattavat realisoitua vasta pidemmällä aikavälillä (Jansen & Stroop, 2001). Huomioon on otettava myös se tosiasia, että henkilöarviointi on vain yksi osa koko rekrytointiprosessia. Näin ollen esihenkilön antama arvio henkilön työmenestyksestä ei kerro pelkästään Psycon Oy:n raportin osuvuudesta, vaan koko rekrytointiprosessin onnistuneisuudesta. On kuitenkin syytä ajatella, että henkilöarviointipalveluita hyödyntävät organisaatiot nojaavat päätöksenteossaan teetetyn henkilöarvioinnin tuloksiin, sillä juuri tulokset ovat keskeinen henkilöarvioinnin tuottama lisäarvo. Yli 96 % tässä tutkimuksessa tarkastelluista rekrytoinneista oli joko erinomaisesti tai hyvin onnistuneita, mikä tukee ajatusta tämänkaltaisen rekrytointiprosessin toimivuudesta.

Tilanteet, joissa henkilöarviointia voidaan pitää heikosti osuvana, ovat käytännössä ne, joissa suosituksen ja työmenestyksen erotuspistemäärä on 2 tai 3 pistettä. Tällöin henkilöarvioinnin raportin suositus voi esimerkiksi olla paras ja esihenkilöarvio toiseksi huonoin. Suositusta voidaan pitää heikosti osuvana myös silloin, kun se on ollut selvästi liian kriittinen, eli lopulta hyvin menestynyt henkilö on saanut alun perin heikon suosituksen. Näitä esimerkkejä oli aineistossa lukumäärällisesti vähän jo siitä syystä, että rekrytoitujen joukossa on yleensäkin vähän matalamman suosituksen saaneita henkilöitä. Näin ollen aineiston perusteella on vaikea arvioida sitä, onko osa arvioitavista henkilöistä saanut selvästi liian pessimistisen suosituksen.

Tämän tutkimuksen perusteella voidaan todeta, että henkilöarvioinnin suositukset ovat keskimäärin esihenkilöarvioita kriittisempiä. Sama havaittiin myös Heikkilän (2017) tutkimuksessa. Arviointia tehtäessä on tasapainoteltava väärän positiivisen suosituksen ja väärän negatiivisen suosituksen välillä. Kun huomioon otetaan väärästä rekrytoinnista koituvat haitat, kuten taloudelliset menetykset, on ymmärrettävää, että henkilöarvioinnin raporteissa arviot ovat enemmän hieman konservatiivisia kuin ylioptimistisia (Karsan, 2007; Montgomery-Young, 2018). Tästä koituvana vaarana kuitenkin on, että työhön soveltuva henkilö ei saa työpaikkaa ylikriittisen arviointisuosituksen vuoksi. Epäsuorasti haitat voivat näkyä myös yhteiskunnan tasolla, mikäli todellisuudessa hyvät työntekijät jäävät työpaikkojen ulkopuolelle ja kysymys onkin oleellinen myös eettisesti. Kaiken kaikkiaan on kuitenkin perustellumpaa, että työpaikoista joissakin tapauksissa väistämättä käytävä kilpailu ratkaistaan mahdollisimman luotettavilla ja tasapuolisilla keinoilla, kuten laadukkaalla henkilöarvioinnilla.

Koska henkilöarvioinnin tarkoituksena on tuottaa tietoa päätöksenteon tueksi, arviointikonsulteilta odotetaan kannanottoa henkilön soveltuvuuteen (Suomen Psykologiliitto ry, 2019). Etätilanteessa yli puolet rekrytoiduista henkilöistä sai suositukseen parhaan mahdollisen arvion ja läsnätilanteessa parhaan suosituksen saaneita oli hieman alle puolet. Lukujen pohjalta voidaan todeta, että vaikka henkilöarviointien suositukset ovat kriittisempiä kuin esihenkilöiden arviot, niissä myös otetaan kantaa henkilön soveltuvuuteen ja uskalletaan suositella. Paras henkilöarviointiraportin suositus tarkoittaa ”hän täyttää tehtävän asettamat vaatimukset ja valintaa voi suositella” ja toiseksi paras ”hän täyttää tehtävän asettamat vaatimukset suurelta osin, joskin panostaminen kehittämisalueisiin parantaisi menestysennustetta”. Kokonaisuudesta voidaan todeta, että kaikkien kriteerien täyttyminen ei rekrytoinnissa aina ole itsestään selvää, mutta kehitysalueet eivät välttämättä estä työssä menestymistä. Lisäksi on syytä huomioda, että rekrytoitujen henkilöiden kehitysalueiden tunnistaminen jo rekrytointivaiheessa voi toimia henkilön kehityksen tukena jatkossa. On

mahdollista, että osa henkilöarvointiraporttien kriittisyydestä suhteessa esihenkilöiden arvioihin selittyy sillä, että rekrytoidun henkilön kehityskohteet on tunnistettu ajoissa ja henkilön työssä menestymistä on osattu tukea paremmin. Näin ollen ero arvioissa voi itseasiassa viestiä myös raportin hyödyllisyydestä osana rekrytointia ja perehdytystä. Myös Honkanen (2005) mainitsee, että henkilöarvointia voi hyvin hyödyntää osana henkilöstön kehittämistä.

4.2 Yksittäisten menetelmien ennustavuus

Yksittäiset menetelmät ennustivat työmenestystä melko heikosti, sillä koko aineiston tasolla tarkasteltuna yksikään yksittäinen psykologinen ominaisuus ei kyennyt selittämään yli 10 % kokonaisvaihtelusta. Teorian pohjalta oletettiin, että kognitiivinen kyvykkyys ja henkilön tunnollisuus ovat yhteydessä työmenestykseen (Arthur ym., 2003; Barrick & Mount, 1991; Bobko ym., 1999; Hunter & Hunter, 1984; Morris ym., 2015; Schmidt ym., 2008; Schmidt & Hunter, 1998). Tunnollisuus ei kuitenkaan ennustanut työmenestystä tässä tutkimuksessa. Myöskään Heikkilän (2017) tutkimuksessa Psycon Oy:n henkilöarvioinnin ennustavuudesta ei havaittu tunnollisuuden olevan yhteydessä työmenestykseen. Tunnollisuus ei siis vaikuta erottelevan hakijoita ainakaan Psycon Oy:n arvioimien rekrytoitujen henkilöiden populaatiossa. Sen sijaan numeerisinduktiivinen päättely oli yhteydessä menestymiseen tässä tutkimuksessa tarkastelluista menetelmistä koko aineistossa. Lisäksi induktiivinen päättely ennusti työmenestystä läsnätilanteessa. Näin ollen voidaan todeta, että kognitiivinen kyvykkyys ennustaa työmenestystä myös tämän aineiston perusteella, joskaan efekti ei ollut kovinkaan suuri. Tulos on yhdenmukainen aikaisempien tutkimusten kanssa, joissa kognitiivisen kyvykkyuden on havaittu ennustavan työmenestystä (Arthur ym., 2003; Bobko ym., 1999; Hunter & Hunter, 1984; Morris ym., 2015; Schmidt ym., 2008; Schmidt & Hunter, 1998). Erityisen suuri merkitys vaikuttaa olevan kyvyllä tehdä päätelmiä ja hahmottaa kokonaisuuksia.

Vaikka tunnollisuuden on aikaisemmin havaittu olevan yhteydessä työmenestykseen, esimerkiksi Huang (2014) argumentoi tunnollisuuden olevan keskinkertainen työmenestyksen ennustaja. Piirteen eduksi hän mainitsee sen, että tunnollisuuden kyky ennustaa työmenestystä säilyy melko samanlaisena työnkuvasta ja työn kriteereistä riippumatta. Kun huomioon otetaan lisäksi havainto siitä, että persoonallisuustestien kyky ennustaa työmenestystä yleisesti on varsin heikko, tämän tutkimuksen tulos ei ole erityisen dramaattinen (Huang ym., 2014; Hurtz & Donovan, 2000; Morgeson ym., 2007; Morris ym., 2015; Ones & Viswesvaran, 1996;

Schmidt ym., 2008). On myös mahdollista, että nykyään työelämässä korostuvat tunnollisuuden sijasta muut ominaisuudet.

Tämän tutkimuksen osalta tunnollisuuden heikkoa efektiä voidaan selittää myös sillä, että todennäköisesti henkilöt, joiden kohdalla liian matala tai liian korkea tunnollisuus olisi ollut merkittävä riski työssä menestymiselle, on tunnistettu viimeistään arviointivaiheessa. Tämän karsiutumisen vuoksi piirteen kyky ennustaa työmenestystä voi jäädä heikoksi rekrytoitujen keskuudessa. Sama koskee jokaista tässä tutkimuksessa tarkasteltua yksittäistä menetelmää ja niiden dimensiota, ja on mahdollista, että osa vaisuista efekteistä selittyy juuri valikoitumisella. Se, että yksittäiset ominaisuudet eivät enää selitä varianssia rekrytoitujen henkilöiden keskuudessa, voi toisin sanoen olla osoitus siitä, että työmenestyksen kannalta keskeiset piirteet on tunnistettu henkilöarvioinnin yhteydessä, ne on huomioitu suosituksessa, eivätkä ne sen vuoksi selitä vaihtelua enää valituksi tulleiden henkilöiden keskuudessa.

Edellisen lisäksi arviointiin osallistuvat henkilöt ovat yleensä jo lähtökohtaisesti valikoitunut joukko monien ominaisuuksien suhteen (Dahlke & Wiernik, 2020, Lievens & Thornton, 2005). Tähän vaikuttavat sekä rekrytoinnin alkuseulonta ennen arvioinnin teettämistä, että se, että henkilöarviointeja saatetaan hyödyntää organisaatioissa vain korkeamman tason rekrytoinneissa melko korkeiden kustannusten vuoksi. Tässäkin tutkimuksessa ylivoimainen enemmistö arvioinneista oli asiantuntija- esimies- ja johtotehtäviin. Tärkeää on huomioda myös havainto siitä, että persoonallisuuden osalta tulee tarkastella laajempia profiileja sekä henkilön kognitiivista kyvykkyyttä ja muita ominaisuuksia osana kokonaisuutta, sillä eri tekijöillä voi olla toisiinsa meroeroivia vaikutuksia (Conte ym., 2017; Perkins & Corr, 2006). Tämän tutkimuksen pohjalta ei siis voida vielä päätellä, että persoonallisuus on lähtökohtaisesti epäolennainen osa-alue työmenestyksen kannalta.

Tutkimuksessa tarkasteltiin yhteensä 28 persoonallisuuden ulottuvuuden yhteyttä työmenestykseen, mutta vain kaksi ulottuvuutta oli työmenestykseen yhteydessä. Persoonallisuuden ulottuvuuksista estyneisyys heikensi työmenestyksen ennustetta sekä tässä että Heikkilän (2017) tutkimuksessa. Estyneisyys liittyy sosiaaliseen pelkoon, huomion välttämiseen ja arkuuteen uusien ihmisten tapaamisessa (Psycon Oy, 2017c). Piirteeseen liittyy myös kritiikin ja epäonnistumisen pelko, mikä voi johtaa välttämiskäyttäytymiseen ja liialliseen kontrolloinnin tarpeeseen. Vaikka estyneisyyteen voi toisaalta liittyä tunnollisuutta ja positiiviseksi katsottua itsehillintää, voi liiallinen arkuus estää esimerkiksi sosiaalista

lähentymistä ja toimeen tarttumista, mikä puolestaan heijastuu negatiivisesti työmenestykseen. Juuri estyneisyyden negatiiviset vaikutukset näyttävät korostuvan työmenestyksen osalta.

Myös defensiivisyys oli negatiivisesti yhteydessä työmenestykseen. Defensiivisyys viittaa passiivis-aggressiivisuuteen, vainoharhaisuuteen ja piilotettuun hyökkäävyyteen (Psycon Oy, 2017b). Defensiivisyyteen liittyy niin ikään herkkyys kritiikille. Todennäköistä on, että defensiivisyyden negatiivinen vaikutus työmenestykseen liittyy ennen kaikkea piirteen sosiaaliseen ulottuvuuteen ja defensiivisen henkilön mahdolliseen vaikeuteen luoda rakentavia ihmissuhteita. Monella työpaikalla kyky sujuvaan kommunikaatioon ymmärrettävästi korostuu myös työssä menestymisen kannalta. Lisäksi defensiivisyys voi heikentää henkilön kykyä vastaanottaa palautetta ja kehittyä sen avulla.

Defensiivisyyden ja estyneisyyden esiin nousemisesta huolimatta kokonaisuudesta voidaan todeta, että persoonallisuuden piirteet erottelivat hakijoita heikosti ja niiden yhteys työmenestykseen oli vaatimaton rekrytoitujen keskuudessa. Tämä tukee aiemmin esitettyä ajatusta siitä, että yksittäisten persoonallisuuden piirteiden merkitystä osana rekrytointia ei lähtökohtaisesti pidä korostaa. Tutkimuksessa persoonallisuuden ulottuvuuksien kyky ennustaa työmenestystä ei ollut riippuvainen siitä, oliko kyseessä etätilanne vai läsnätilanne. Näin ollen etäaikakaudella yksittäiset persoonallisuuden piirteet eivät vaikuta selkeällä tavalla korostuvan verrattuna aiempaan.

4.3 Tutkimuksen vahvuudet ja rajoitteet

Tämä tutkimus tuotti uutta tietoa etäyhteydellä toteutetun henkilöarvioinnin ennustavuudesta ja laajensi henkilöarvioinnin ennustavuutta käsittelevää tutkimuskirjallisuutta. Tutkimus tuotti tietoa myös etäarvioinnin ennustavuudesta verrattuna perinteiseen henkilöarviointiin. Tämän kaltaista tutkimusta ei tiettävästi ole tehty koskaan aikaisemmin ja tutkimuksen tuloksilla voi olla keskeinen käytännön merkitys henkilöarviointien toteutuksessa tästä eteenpäin. Lisäksi tutkimuksen aineistot etätilanteesta ja läsnätilanteesta olivat keskenään verrannollista niin käytettyjen menetelmien kuin otoksen suhteen. Yhtenä tutkimuksen etuna voidaan pitää myös sitä, että sen otanta koostui aidon henkilöarviointitilanteen läpikäyneistä rekrytoiduista henkilöistä ja näin ollen tutkimuksen tuloksia voidaan pitää näiltä osin edustavana kohdepopulaatiossa.

Tutkimuksen yhtenä vahvuutena ovat myös Psycon Oy:n käyttämät validit ja reliabelit arviointimenetelmät, joiden voidaan todeta mittaavan niitä ominaisuuksia, joita niiden

kerrotaan mittaavan. Yritys myös käyttää omia menetelmiä, minkä ansiosta arviointikonsultit saavat kattavan perehdytyksen ja laajaa tukea menetelmien soveltamiseen. Lisäksi arvioitavia persoonallisuuden, toimintatyylin ja kognitiivisen kyvykkyyden dimensioita oli tässä tutkimuksessa yhteensä yli 30, minkä ansiosta tutkimuksessa saatiin tietoa laaja-alaisesti persoonallisuuden ja kognitiivisen kyvykkyyden kyvystä ennustaa työmenestystä. Myös tutkimuksen otoskokoa voidaan pitää kohtuullisen suurena. Esihenkilöarvioiden käyttäminen työmenestyksen mittarina on vakiintunut käytäntö ja näin ollen tulosten vertaaminen aikaisempaan tutkimukseen oli mielekästä.

Etukäteen osattiin ennakoida, että vaihteluvälin kapeutuminen näkyy aineistossa, sillä ongelma on tyypillinen tämän kaltaiselle tutkimukselle (Dahlke & Wiernik, 2020, Lievens & Thornton, 2005). Henkilöarvioinnista eteenpäin jatkavat tyypillisesti hyvän suosituksen saaneet ja jo ennen arviointia rekrytoitavat ovat läpikäyneet tiukan seulan. Lisäksi henkilöarvioita käytetään tavanomaisesti vain tietyssä osassa rekrytointeja, jolloin valikoitumista tapahtuu myös työnkuvien suhteen. Valikoitumisen myötä kriittisiä arvioita esiintyi tutkimusaineistossa vain vähän. Tässä tutkimuksessa tutkimusasetelma säilytettiin valikoitumisesta huolimatta samankaltaisena kuin aikaisemmissa Psycon Oy:n tutkimuksissa, jotta aineistojen vertailu etä- ja läsnätilanteesta oli mahdollista. Puutteistaan huolimatta esihenkilöiden haastattelemisen työhön valituista henkilöistä on menettely, joka mahdollistaa tiedon keräämisen henkilöarvioinnin ennustavuudesta ja henkilöiden työmenestyksestä.

Tässä tutkimuksessa valikoitumisesta aiheutuneeseen vaihteluvälin kapeutumiseen reagoitiin käyttämällä epäparametrista khiin neliö -testiä sekä dikotomiseen jaotteluun perustuvaa logistista regressiota. Logistisen regression käyttäminen kuitenkin tarkoitti sitä, että todellisuudessa luonteeltaan jatkuvaa informaatiota joutui tiivistämään dikotomiseen muotoon ja näin ollen datan informatiivisuus saattoi kärsiä. Tilastomenetelmien käyttöön liittyvien huomioiden lisäksi yksi vaihteluvälin kapeutumisesta aiheutuva ongelma on, että liian kriittisistä henkilöarvioinnin suosituksista ja niiden vaikutuksista kertyy tällaisessa tutkimuksessa heikosti tietoa. Tutkimuksen rajoituksena voidaan pitää myös sen heikkoa yleistettävyyttä normaalipopulaatioon, sillä vain osassa rekrytointeja käytetään henkilöarviointia tässä laajuudessa.

Tämänkaltaisessa tutkimuksessa on käytännön tasolla erittäin vaikeaa luoda asetelma, jossa työmenestystä tarkasteltaisiin myös verrokkiryhmän osalta. Verrokkiryhmä voisi koostua esimerkiksi pelkät kognitiiviset kykytehtävät suorittaneista henkilöistä. Verrokkiryhmän

puuttumisen vuoksi ei voitu tehdä vahvoja oletuksia syy-seuraussuhteista. Tässä tutkimuksessa mukaan ei otettu assistenttihakintoja tai simulaatioissa suoriutumista, mitä voidaan pitää tutkimuksen rajoituksena, sillä niiden on aikaisemmissa tutkimuksissa havaittu olevan yhteydessä työmenestykseen (Hermelin ym., 2007; Lievens & Patterson, 2011; O'Connel ym., 2002). Myöskään työn kriteereitä ei kontrolloitu, minkä vuoksi sattuman roolia esimerkiksi yksittäisten menetelmien ennustavuuden suhteen on vaikea arvioida. Työn kriteereillä viitataan niihin työn ominaisuuksiin, minkä pohjalta arvio soveltuvuudesta tehdään.

Tutkimuksen otokseen liittyy myös se ongelma, että kaksi henkilöä, jotka omasta aloitteestaan päättivät työsuhteensa koeajalla, jäivät analyysien ulkopuolelle. Näiden tapausten näkeminen automaattisesti epäonnistuneina henkilöarviointina ei ole perusteltua, sillä rekrytoivat osapuolet ovat voineet olla tietoisia henkilön motivaatioon liittyneistä riskeistä, eikä työsuhteen päättäminen myöskään tarkoita heikkoa soveltuvuutta. Osaajista kilpailemista pidetään jopa tyypillisenä piirteenä tämän päivän työelämälle, jolloin myös työpaikan vaihtaminen voidaan nähdä luonnollisena osana työelämää. Tutkimuksen otoskoko kärsi myös puuttuvien havaintojen vuoksi.

Vaikka etähenkilöarviointien ennustavuuden samankaltaisuus lähikontaktissa toteutettujen henkilöarvioiden kanssa viittaa siihen, että huijaaminen ei ole iso ongelma etäarvioinneissa, ei huijaamisen mahdollisuutta voi sulkea täysin pois. Mikäli esimerkiksi kognitiiviset kykytehtävät on tehnyt arvioitavan henkilön puolesta joku muu, vääristää se myös tässä tutkimuksessa saatuja tuloksia yksittäisten menetelmien ennustavuudesta. Huolta huijaamisesta on pidetty erityisen suurena juuri kognitiivisten kykytehtävien suhteen (Tippins ym., 2006). Yhtenä tutkimuksen rajoituksena voidaan pitää myös sitä, että analyysit olivat pelkästään kvantitatiivisia. Eri osapuolien henkilökohtaiset kokemukset ja laadullinen informaatio prosessin toimivuudesta jäivät tämän tutkimuksen ulkopuolelle, vaikka ne olisivat voineet antaa hyödyllistä lisätietoa. Tämän tutkimuksen perusteella voidaan todeta, että myös etähenkilöarviointi kykenee ennustamaan työmenestystä, mutta esimerkiksi kokemuksen mielekkyydestä ei voida tehdä johtopäätöksiä.

Yksi tämän tutkimuksen rajoituksista liittyy siihen, että Psycon Oy:n lisäksi myös valtaosa asiakasyrityksistä on toiminut etäyhteyksien varassa. Näin ollen etähenkilöarvioinnin voidaan tämän tutkimuksen perusteella todeta olevan toimiva malli varmasti vain silloin, kun myös suoritettava työ tapahtuu pääsääntöisesti etäyhteydellä. Koska työn kriteereitä ei kontrolloitu, myöskään etätyön konkreettisista vaikutuksista työnkuviin ei voida tehdä voimakkaita

johtopäätöksiä. Howlandin (2015) hypoteesi siitä, että arviointivälineen mukauttaminen varsinaista työtehtävää vastaavaksi parantaa ennustetarkkuutta, ei tämän tutkimuksen perusteella ole ainakaan pois suljettu vaihtoehto. Huomioitavaa on myös se, että työhön rekrytoitujen henkilöiden perehdytys on tapahtunut korona-aikana usein etäyhteydellä ja lähikontaktit esihenkilöiden kanssa ovat olleet vähäisempiä. Tämä olisi mahdollisesti voinut vaikuttaa siten, että esihenkilöt ovat arvioinneissaan varovaisempia ja antavat täten vähemmän kriittisiä arvioita. Tulosten perusteella esihenkilöt kuitenkin kokivat olevansa yhtä varmoja omista arvioistaan etätilanteessa kuin läsnätilanteessa ja lisäksi arvioinnin epävarmuus oli yhteydessä etenkin varovaisempiin positiivisiin arvioihin.

Tutkimuksen rajoitteeksi voidaan mainita myös se, että henkilön työmenestyksen mittarina käytettiin ainoastaan esihenkilön arviota, eikä useita rinnakkaisia menetelmiä, kuten itsearviointeja ja toveriarvioita tai numeerisia mittareita esimerkiksi palkkakehityksestä. Esihenkilöarviointien arvioijien väliset (inter-rater reliability) reliabiliteetit ovat vain noin .50, mikä osaltaan korostaa pelkästään kyseisen mittarin käytön rajoittuneisuutta (Viswesvaran ym., 1996). Useiden mittareiden käyttäminen olisi toisaalta varsin raskasta asiakasyritysten kannalta, jolloin riittävän suuren otoksen kokoaminen voisi muotoutua ongelmaksi. Tässä tutkimuksessa pelkällä esihenkilöarvioinnilla työmenestyksestä ei kuitenkaan saatu esiin merkittäviä eroja rekrytoitujen henkilöiden välillä.

Lopuksi maininnan arvoista on se, että henkilöarvioinnin yleisarvio on henkilön soveltuvuus tiivistettynä yhteen numeroon. Käytännössä on hyvin mahdollista, että useampi kandidaatti saa henkilöarvioinnista saman suosituksen, jolloin päätöksenteossa korostuvat yksittäiset ominaisuudet ja organisaation eniten painottamat kriteerit. Tätä hienovaraisempaa informaatiota ei kuitenkaan ollut mahdollista selvittää tässä tutkimuksessa. Kaiken kaikkiaan tutkimuksen rajoituksista heijastuu hyvin se, kuinka vaativaa henkilöarvioinnin osuvuutta on lopulta tutkia.

4.4 Jatkotutkimusehdotukset

Jatkotutkimuksen osalta tärkeä alue on ensinnäkin etähenkilöarvioinnin ennustavuuden tutkiminen tilanteessa, jossa pandemia ei välittömästi vaikuta ihmisten työskentelyyn ja jolloin asiakasyritysten toimintatavat eivät todennäköisesti painotu ainakaan aivan yhtä vahvasti etätyöhön. Toiseksi tämän tutkimuksen aineistossa henkilöarviointiraporttien liian kriittisistä suosituksista saatiin niukasti tietoa vaihteluvälin kapeutumisen vuoksi. Tästä muutenkin

tyypillisesti katveeseen jäävästä tutkimusalueesta on hyvä saada tulevaisuudessa lisää tutkittua tietoa, jotta tietoa pystytään hyödyntämään myös henkilöarviointiprosessien kehittämisessä.

Kolmanneksi jatkossa olisi hyvä tutkia simulaatioiden ja muiden vuorovaikutuksellisten menetelmien kykyä ennustaa työmenestystä etähenkilöarvioinnissa, sillä näiden menetelmien on aikaisemmassa henkilöarviointia koskevassa tutkimuksessa havaittu olevan yhteydessä työmenestykseen (Hermelin ym., 2007; Lievens & Patterson, 2011; O’Connel ym., 2002). Neljänneksi olisi tärkeää, että jatkossa etenkin etähenkilöarvioinnin ennustavuudesta tehtäisiin laajoja tutkimuksia, joissa käytetään rinnakkain useampia työmenestyksen mittareita, jolloin varianssin työmenestyksessä voitaisiin olettaa lisääntyvän. Varianssi puolestaan on tärkeää tilastollisten analyysien ja niistä tehtävien johtopäätösten kannalta. Myös henkilöarvioinnin yleisarvion rinnalla voitaisiin jatkossa käyttää useampia mittareita, kuten esihenkilöarvioita, jolloin olisi helpompi tehdä johtopäätöksiä juuri henkilöarvioinnin yleisarvion hyödyllisyydestä osana rekrytointia. Lisäksi työmenestyksen tarkasteleminen myös pidemmällä aikavälillä on perusteltua, sillä osa työmenestykseen vaikuttavista piirteistä voi tulla esiin vasta ajan kuluessa.

Vaikka lisää tutkimusta tarvitaan, voi jo tämän tutkimuksen pohjalta todeta, että henkilöarvioinnilla kyetään ennustamaan työmenestystä niin etäyhteydellä kuin lähikontaktissa toteutettuna. Näin ollen laadukas rekrytointiprosessi voidaan toteuttaa yhä joustavammin ja sekä rekrytoivan yrityksen että työnhakijan tarpeet huomioiden. Tämä tutkimus on yksi osoitus teknologian luomista mahdollisuuksista psykologian kaltaisella alalla, jossa keskiössä on lopulta ennen kaikkea ihminen.

Lähteet:

- Arthur, W., Jr., Day, E. A., McNelly, T. L., & Edens, P. S. (2003). A meta-analysis of the criterion-related validity of assessment center dimensions. *Personnel Psychology*, 56(1), 125–154. <https://doi-org.ezproxy.utu.fi/10.1111/j.1744-6570.2003.tb00146.x>.
- Barak, A., & Buchanan, T. (2004). Internet-based psychological testing and assessment. Teoksessa Kraus, R., & Zack, J. *Online Counseling: A Handbook for Mental Health Professionals* (s. 217–239). Elsevier.
- Barak, A., & English, N. (2002). Prospects and limitations of psychological testing on the Internet. *Journal of Technology in Human Services*, 19(2–3), 65–89. https://doi.org/10.1300/J017v19n02_06.
- Barrick, M. R., Dustin, S. L., Giluk, T. L., Stewart, G. L., Shaffer, J. A., & Swider, B. W. (2012). Candidate characteristics driving initial impressions during rapport building: Implications for employment interview validity. *Journal of Occupational and Organizational Psychology*, 85(2), 330–352. <https://doi.org/10.1111/j.2044-8325.2011.02036.x>.
- Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, 44(1), 1–26. <https://doi-org.ezproxy.utu.fi/10.1111/j.1744-6570.1991.tb00688.x>.
- Barrick, M. R., Shaffer, J. A., & DeGrassi, S. W. (2009). What you see may not be what you get: Relationships among self-presentation tactics and ratings of interview and job performance. *Journal of Applied Psychology*, 94(6), 1394–1411. <https://doi-org.ezproxy.utu.fi/10.1037/a0016532>.
- Bartels, L. K., Bommer, W. H., & Rubin, R. S. (2000). Student performance: Assessment centers versus traditional classroom evaluation techniques. *Journal of Education for Business*, 75(4), 198–201. <https://doi.org/10.1080/08832320009599014>.
- Bartram, D. (1997). Distance assessment: Psychological assessment through the Internet. *Selection and Development Review*, 13(3), 15–19.
- Bartram, D. (1999). Testing and the Internet: Current realities, issues and future possibilities. *Selection and Development Review*, 15(6), 3–12.
- Basch, J., Melchers, K., Kurz, A., Krieger, M., & Miller, L. (2020). It takes more than a good camera: Which factors contribute to differences between face-to-face Interviews and videoconference interviews regarding performance ratings and interviewee

- perceptions? *Journal of Business and Psychology*, 1–20. <https://doi.org/10.1007/s10869-020-09714-3>
- Blackman, M. C. (2002). The employment interview via the telephone: Are we sacrificing accurate personality judgments for cost efficiency? *Journal of Research in Personality*, 36(3), 208–223. <https://doi-org.ezproxy.utu.fi/10.1006/jrpe.2001.2347>.
- Blacksmith, N., Willford, J. C., & Behrend, T. S. (2016). Technology in the employment interview: A meta-analysis and future research agenda. *Personnel Assessment and Decisions*, 2(1), 2. <https://doi.org/10.25035/pad.2016.002>
- Bobko, P., Roth, P. L., & Potosky, D. (1999). Derivation and implications of a meta-analytic matrix incorporating cognitive ability, alternative predictors, and job performance. *Personnel Psychology*, 52(3), 561–589. <https://doi-org.ezproxy.utu.fi/10.1111/j.1744-6570.1999.tb00172.x>.
- Buchanan, T. (2003). Internet-based questionnaire assessment: Appropriate use in clinical contexts. *Cognitive Behaviour Therapy*, 32(3), 100–109. <https://doi-org.ezproxy.utu.fi/10.1080/16506070310000957>.
- Campos, J. A. D. B., Zucoloto, M. L., Bonafé, F. S. S., Jordani, P. C., & Maroco, J. (2011). Reliability and validity of self-reported burnout in college students: A cross randomized comparison of paper-and-pencil vs. online administration. *Computers in Human Behavior*, 27(5), 1875–1883.
- Carter, N. T., Dalal, D. K., Boyce, A. S., O'Connell, M. S., Kung, M.-C., & Delgado, K. M. (2014). Uncovering curvilinear relationships between conscientiousness and job performance: How theoretically appropriate measurement makes an empirical difference. *Journal of Applied Psychology*, 99(4), 564–586. <https://doi-org.ezproxy.utu.fi/10.1037/a0034688>.
- Chapman, D. S., & Rowe, P. M. (2001). The impact of videoconference technology, interview structure, and interviewer gender on interviewer evaluations in the employment interview: A field experiment. *Journal of Occupational and Organizational Psychology*, 74(3), 279–298. <https://doi-org.ezproxy.utu.fi/10.1348/096317901167361>.
- Conte, J. M., Heffner, T. S., Roesch, S. C., & Aasen, B. (2017). A person-centric investigation of personality types, job performance, and attrition. *Personality and Individual Differences*, 104, 554–559. <https://doi-org.ezproxy.utu.fi/10.1016/j.paid.2016.09.004>.
- Cronk, B. C., & West, J. L. (2002). Personality research on the internet: A comparison of web-based and traditional instruments in take-home and in-class settings. *Behavior Research*

- Methods, Instruments & Computers*, 34(2), 177–180. <https://doi-org.ezproxy.utu.fi/10.3758/BF03195440>.
- Dahlke, J., & Wiernik, B. (2020). Not restricted to selection research: Accounting for indirect range restriction in organizational research. *Organizational Research Methods*, 23(4), 717–749. <https://doi.org/10.1177/1094428119859398>.
- Damitz, M., Manzey, D., Kleinmann, M., & Severin, K. (2003). Assessment center for pilot selection: Construct and criterion validity and the impact of assessor type. *Applied Psychology: An International Review*, 52(2), 193–212. <https://doi-org.ezproxy.utu.fi/10.1111/1464-0597.00131>.
- Dayan, K., Kasten, R., & Fox, S. (2002). Entry-level police candidate assessment center: An efficient tool or a hammer to kill a fly? *Personnel Psychology*, 55(4), 827–849. <https://doi-org.ezproxy.utu.fi/10.1111/j.1744-6570.2002.tb00131.x>.
- Dudley, N. M., Orvis, K. A., Lebiecki, J. E., & Cortina, J. M. (2006). A meta-analytic investigation of conscientiousness in the prediction of job performance: Examining the intercorrelations and the incremental validity of narrow traits. *Journal of Applied Psychology*, 91(1), 40–57. <https://doi-org.ezproxy.utu.fi/10.1037/0021-9010.91.1.40>.
- Ellis, A. P. J., West, B. J., Ryan, A. M., & DeShon, R. P. (2002). The use of impression management tactics in structured interviews: A function of question type? *Journal of Applied Psychology*, 87(6), 1200–1208. <https://doi-org.ezproxy.utu.fi/10.1037/0021-9010.87.6.1200>.
- Fouladi, R. T., McCarthy, C. J., & Moller, N. P. (2002). Paper-and-pencil or online? Evaluating mode effects on measures of emotional functioning and attachment. *Assessment*, 9(2), 204–215. <https://doi-org.ezproxy.utu.fi/10.1177/10791102009002011>.
- Furnham, A., & Nederstrom, M. (2010). Ability, demographic and personality predictors of creativity. *Personality and Individual Differences*, 48(8), 957–961. <https://doi.org/10.1016/j.paid.2010.02.030>.
- Gati, I., & Saka, N. (2001). Internet-based versus paper-and-pencil assessment: Measuring career decision-making difficulties. *Journal of Career Assessment*, 9(4), 397–416. <https://doi-org.ezproxy.utu.fi/10.1177/106907270100900406>.
- Geddes, M. R., O'Connell, M. E., Fisk, J. D., Gauthier, S., Camicioli, R., & Ismail, Z. (2020). Remote cognitive and behavioral assessment: Report of the Alzheimer Society of Canada Task Force on dementia care best practices for COVID-19. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 12(1), e12111–n/a. <https://doi.org/10.1002/dad2.12111>

- Gore, P. A., & Leuwerke, W. C. (2000). Information technology for career assessment on the Internet. *Journal of Career Assessment*, 8(1), 3–19. <https://doi-org.ezproxy.utu.fi/10.1177/106907270000800102>.
- Heikkilä, R. (2017). Tapaustutkimus: Psycon Oy:n Metsä Groupille vuosina 2014-2016 tekemien henkilöarviointien ennustavuus ja työmenestystä Metsä Groupilla ennustavat psykologiset ominaisuudet.
- Hermelin, E., Lievens, F., & Robertson, I. T. (2007). The validity of assessment centres for the prediction of supervisory performance ratings: A meta-analysis. *International Journal of Selection and Assessment*, 15(4), 405–411. <https://doi-org.ezproxy.utu.fi/10.1111/j.1468-2389.2007.00399.x>.
- Holzman, T. G., Pellegrino, J. W., & Glaser, R. (1983). Cognitive variables in series completion. *Journal of Educational Psychology*, 75(4), 603–618. <https://doi-org.ezproxy.utu.fi/10.1037/0022-0663.75.4.603>.
- Honkanen, H. (2005). *Henkilöarviointi työelämässä*. Edita.
- Horn, R., & Behrend, T. (2017). Video killed the interview star: Does picture-in-picture affect interview performance? *Personnel Assessment and Decisions*, 3(1). <https://doi.org/10.25035/pad.2017.005>.
- Howland, A. C., Rembisz, R., Wang-Jones, T. S., Heise, S. R., & Brown, S. (2015). Developing a virtual assessment center. *Consulting Psychology Journal: Practice and Research*, 67(2), 110. <https://doi.org/10.1037/cpb0000034>.
- Huang, J. L., Ryan, A. M., Zabel, K. L., & Palmer, A. (2014). Personality and adaptive performance at work: A meta-analytic investigation. *Journal of Applied Psychology*, 99(1), 162–179. <https://doi-org.ezproxy.utu.fi/10.1037/a0034285>.
- Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin* 96(1), 72–98. <https://doi-org.ezproxy.utu.fi/10.1037/0033-2909.96.1.72>.
- Hurtz, G. M., & Donovan, J. J. (2000). Personality and job performance: The Big Five revisited. *Journal of Applied Psychology*, 85(6), 869–879. <https://doi-org.ezproxy.utu.fi/10.1037/0021-9010.85.6.869>.
- Ingold, P. V., Kleinmann, M., König, C. J., & Melchers, K. G. (2015). Shall we continue or stop disapproving of self-presentation? Evidence on impression management and faking in a selection context and their relation to job performance. *European Journal of Work and Organizational Psychology*, 24(3), 420–432. <https://doi-org.ezproxy.utu.fi/10.1080/1359432X.2014.915215>.

- Jansen, P. G. W., & Stoop, B. A. M. (2001). The dynamics of assessment center validity: Results of a 7-year study. *Journal of Applied Psychology*, 86(4), 741–753. <https://doi-org.ezproxy.utu.fi/10.1037/0021-9010.86.4.741>.
- Joinson, A. (1999). Social desirability, anonymity, and Internet-based questionnaires. *Behavior Research Methods, Instruments, & Computers*, 31(3), 433–438. <https://doi.org/10.3758/BF03200723>.
- Jones, M., & Sugden, R. (2001). Positive confirmation bias in the acquisition of information. *Theory and Decision*, 50(1), 59–99. <https://doi.org/10.1023/A:1005296023424>.
- Jones, J. W., Brasher, E. E., & Huff, J. W. (2002). Innovations in integrity-based personnel selection: Building a technology-friendly assessment. *International Journal of Selection and Assessment*, 10(1–2), 87–97. <https://doi-org.ezproxy.utu.fi/10.1111/1468-2389.00195>.
- Kaplan, Robert M., and Dennis P. Saccuzzo. (2018). *Psychological Testing: Principles, Applications, and Issues*. Ninth edition. Cengage Learning.
- Karsan, R. (2007). Calculating the cost of turnover. *Employment Relations Today*, 34(1), 33–36.
- Klaczynski, P. A., Gordon, D. H., & Fauth, J. (1997). Goal-oriented critical reasoning and individual differences in critical reasoning biases. *Journal of Educational Psychology*, 89(3), 470–485. <https://doi-org.ezproxy.utu.fi/10.1037/0022-0663.89.3.470>.
- Klimoski, R., & Brickner, M. (1987). Why do assessment centers work? The puzzle of assessment center validity. *Personnel Psychology*, 40(2), 243–260. <https://doi-org.ezproxy.utu.fi/10.1111/j.1744-6570.1987.tb00603.x>.
- Kohonen, I., Kuula-Luumi, A., & Spoof, S. K. (2019). Ihmiseen kohdistuvan tutkimuksen eettiset periaatteet ja ihmistieteiden eettinen ennakkoarviointi Suomessa. *Tutkimuseettisen neuvottelukunnan ohje*, 2021-01.
- Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, 110(15), 5802–5805. <https://doi-org.ezproxy.utu.fi/10.1073/pnas.1218772110>.
- Le, H., Oh, I.-S., Robbins, S. B., Ilies, R., Holland, E., & Westrick, P. (2011). Too much of a good thing: Curvilinear relationships between personality traits and job performance. *Journal of Applied Psychology*, 96(1), 113–133. <https://doi-org.ezproxy.utu.fi/10.1037/a0021016>.

- Lievens, F., Harris, M. M., Van Keer, E., & Bisqueret, C. (2003). Predicting cross-cultural training performance: The validity of personality, cognitive ability, and dimensions measured by an assessment center and a behavior description interview. *Journal of Applied Psychology*, 88(3), 476–489. <https://doi-org.ezproxy.utu.fi/10.1037/0021-9010.88.3.476>.
- Lievens, F., & Patterson, F. (2011). The validity and incremental validity of knowledge tests, low-fidelity simulations, and high-fidelity simulations for predicting job performance in advanced-level high-stakes selection. *Journal of Applied Psychology*, 96(5), 927–940. <https://doi-org.ezproxy.utu.fi/10.1037/a0023496>.
- Lievens, F., & Thornton III, G. C. (2005). Assessment centers: Recent developments in practice and research. *The Blackwell Handbook of Personnel Selection*, 243–264. <https://doi.org/10.1002/9781405164221.ch11>.
- Lin, W.-Y., Zhang, X., Song, H., & Omori, K. (2016). Health information seeking in the Web 2.0 age: Trust in social media, uncertainty reduction, and self-disclosure. *Computers in Human Behavior*, 56, 289–294. <https://doi-org.ezproxy.utu.fi/10.1016/j.chb.2015.11.055>.
- Meade, A., Michels, L., & Lautenschlager, G. (2007). Are internet and paper-and-pencil personality tests truly comparable? An experimental design measurement invariance study. *Organizational Research Methods*, 10(2), 322–345. <https://doi.org/10.1177/1094428106289393>.
- Meriac, J., Hoffman, B., Woehr, D., & Fleisher, M. (2008). Further evidence for the validity of Assessment Center dimensions: A meta-analysis of the incremental criterion-related validity of dimension ratings. *Journal of Applied Psychology*, 93(5), 1042–1052. <https://doi.org/10.1037/0021-9010.93.5.1042>.
- Montgomery-Young, R. (3.1.2018). The ROI of recruitment and onboarding. *HRD-Connect*. Lainattu 15.2.2021, saatavilla: <https://www.hrdconnect.com/2018/01/03/the-roi-of-recruitment-and-onboarding-2/>.
- Mooney, J. (2002). Pre-employment testing on the Internet: Put candidates a click away and hire at modem speed. *Public Personnel Management*, 31(1), 41–52. <https://doi.org/10.1177/009102600203100105>.
- Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K., & Schmitt, N. (2007). Reconsidering the use of personality tests in personnel selection contexts. *Personnel Psychology*, 60(3), 683–729. <https://doi-org.ezproxy.utu.fi/10.1111/j.1744-6570.2007.00089.x>.

- Morris, S. B., Daisley, R. L., Wheeler, M., & Boyer, P. (2015). A meta-analysis of the relationship between individual assessments and job performance. *Journal of Applied Psychology, 100*(1), 5–20. <https://doi-org.ezproxy.utu.fi/10.1037/a0036938.suppl>.
- Nederström, M. (2007). SRS-testin käsikirja. Psycon Oy.
- Nederström, M. (2008). Motivaatiomittarin käsikirja ja tilastolliset liitteet. Psycon Oy.
- Nederström, M., & Furnham, A. (2012). The relationship between the FFM and personality disorders in a personnel selection sample. *Scandinavian Journal of Psychology, 53*(5), 421–429. <https://doi.org/10.1111/j.1467-9450.2012.00964.x>.
- Niitamo, P. (1997). PRF personality research form, käsikirja (manual).
- Nye, C. D., Do, B.-R., Drasgow, F., & Fine, S. (2008). Two-step testing in employee selection: Is score inflation a problem? *International Journal of Selection and Assessment, 16*(2), 112–120. <https://doi-org.ezproxy.utu.fi/10.1111/j.1468-2389.2008.00416.x>.
- Nyman, K., & Honkanen, H. (2001). Hyvän henkilöarvioinnin käsikirja. Psykologien kustannus.
- O’Connell, M. S., Hattrup, K., Doverspike, D., & Cober, A. (2002). The validity of “mini” simulations for Mexican retail salespeople. *Journal of Business and Psychology, 16*(4), 593–600. <https://doi-org.ezproxy.utu.fi/10.1023/A:1015406420028>.
- Oliver, L. W., & Chartrand, J. M. (2000). Strategies for career assessment research on the internet. *Journal of Career Assessment, 8*(1), 95–103. <https://doi-org.ezproxy.utu.fi/10.1177/106907270000800108>.
- Ones, D. S., & Viswesvaran, C. (1996). Bandwidth-fidelity dilemma in personality measurement for personnel selection. *Journal of Organizational Behavior, 17*(6), 609–626. [https://doi-org.ezproxy.utu.fi/10.1002/\(SICI\)1099-1379\(199611\)17:6<609::AID-JOB1828>3.0.CO;2-K](https://doi-org.ezproxy.utu.fi/10.1002/(SICI)1099-1379(199611)17:6<609::AID-JOB1828>3.0.CO;2-K)
- Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology, 46*(3), 598–609. <https://doi-org.ezproxy.utu.fi/10.1037/0022-3514.46.3.598>.
- Perkins, A., & Corr, P. (2006). Cognitive ability as a buffer to neuroticism: Churchill’s secret weapon? *Personality and Individual Differences, 40*(1), 39–51. <https://doi.org/10.1016/j.paid.2005.05.012>.
- Puttonen, M. (2016). Professori tyrmää persoonallisuustestien käytön rekrytoinnissa: ”Johtajuutta ennustavia piirteitä on nolla”. *Helsingin Sanomat* 14.9.2016. Tiede. Lainattu 26.1.2021, saatavilla: <http://www.hs.fi/tiede/a1473738247673>.
- Psycon Oy (2017a). Manuaali Motivaatiomittari 2017.

- Psycon Oy (2017b). Manuaali SRS 2017.
- Psycon Oy (2017c). Psyconin henkilöarviointimenetelmät.
- Psycon Oy (2020). Psyconin arviointimenetelmät.
- Raven, J., JH Court, & Raven, J. C. (1998). Manual for Raven's progressive matrices and vocabulary scales.
- Rentz, D. M., Dekhtyar, M., Sherman, J., Burnham, S., Blacker, D., Aghjayan, S. L., Papp, K. V., Amariglio, R. E., Schembri, A., Chenhall, T., Maruff, P., Aisen, P., Hyman, B. T., & Sperling, R. A. (2016). The feasibility of at-home iPad cognitive testing for use in clinical trials. *The Journal of Prevention of Alzheimer's Disease*, 3(1), 8–12. <https://doi.org/10.14283/jpad.2015.78>.
- Rosenfeld, P. (1997). Impression management, fairness, and the employment interview. *Journal of Business Ethics*, 16(8), 801–808. <https://doi.org/10.1023/A:1017972627516>.
- Rutherford, C., Costa, D., Mercieca-Bebber, R., Rice, H., Gabb, L., & King, M. (2016). Mode of administration does not cause bias in patient-reported outcome results: A meta-analysis. *Quality of Life Research*, 25(3), 559–574. <https://doi.org/10.1007/s11136-015-1110-8>.
- Sackett, P. R., Shewach, O. R., & Keiser, H. N. (2017). Assessment centers versus cognitive ability tests: Challenging the conventional wisdom on criterion-related validity. *Journal of Applied Psychology*, 102(10), 1435–1447. <https://doi-org.ezproxy.utu.fi/10.1037/apl0000236>.
- Sampson, J. (2000). Using the internet to enhance testing in counseling. *Journal of Counseling and Development*, 78(3), 348–356. <https://doi.org/10.1002/j.1556-6676.2000.tb01917.x>.
- Schmitt, N., Gooding, R. Z., Noe, R. A., & Kirsch, M. (1984). Meta-analyses of validity studies published between 1964 and 1982 and the investigation of study characteristics. *Personnel Psychology*, 37(3), 407–422. <https://doi-org.ezproxy.utu.fi/10.1111/j.1744-6570.1984.tb00519.x>.
- Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124(2), 262. <https://doi.org/10.1037//0033-2909.124.2.262>.
- Schmidt, F. L., Shaffer, J. A., & Oh, I.-S. (2008). Increased accuracy for range restriction corrections: Implications for the role of personality and general mental ability in job and training performance. *Personnel Psychology*, 61(4), 827–868. <https://doi-org.ezproxy.utu.fi/10.1111/j.1744-6570.2008.00132.x>.

- Spear, S. E., Shedlin, M., Gilberti, B., Fiellin, M., & McNeely, J. (2016). Feasibility and acceptability of an audio computer-assisted self-interview version of the Alcohol, Smoking and Substance Involvement Screening Test (ASSIST) in primary care patients. *Substance Abuse*, 37(2), 299–305. <https://doi-org.ezproxy.utu.fi/10.1080/08897077.2015.1062460>.
- Stevens, C. K., & Kristof, A. L. (1995). Making the right impression: A field study of applicant impression management during job interviews. *Journal of Applied Psychology*, 80(5), 587–606. <https://doi-org.ezproxy.utu.fi/10.1037/0021-9010.80.5.587>.
- Suomen Psykologiliitto ry. (2019). Henkilöarviointi työelämässä: ohjeistus hyviksi käytännöiksi. Lainattu 26.1.2021, saatavilla: https://www.psyli.fi/tietoa_psykologeista/henkiloarviointi.
- Tapaninen, A., Auer, H., Heiskari, P., Honkaniemi, L., Junnila, K., Kaukonen, T., & Levo, A. (2007). PK5–Persoonallisuustestin Käsikirja. Psykologien Kustannus Oy.
- Tett, R. P., Steele, J. R., & Beauregard, R. S. (2003). Broad and narrow measures on both sides of the personality-job performance relationship. *Journal of Organizational Behavior*, 24(3), 335–356. <https://doi-org.ezproxy.utu.fi/10.1002/job.191>.
- Thornton G. C., & Rupp, D. E. (2006). Assessment centers in human resource management: Strategies for Prediction, Diagnosis, and Development. Lawrence Erlbaum.
- Tippins, N. T. (2009). Internet alternatives to traditional proctored testing: Where are we now? *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 2(1), 2–10. <https://doi-org.ezproxy.utu.fi/10.1111/j.1754-9434.2008.01097.x>.
- Tippins, N. T., Beaty, J., Drasgow, F., Gibson, W. M., Pearlman, K., Segall, D. O., & Shepherd, W. (2006). Unproctored internet testing in employment settings. *Personnel Psychology*, 59(1), 189–225. <https://doi-org.ezproxy.utu.fi/10.1111/j.1744-6570.2006.00909.x>.
- Viswesvaran, C., Ones, D. S., & Schmidt, F. L. (1996). Comparative analysis of the reliability of job performance ratings. *Journal of Applied Psychology*, 81(5), 557–574. <https://doi-org.ezproxy.utu.fi/10.1037/0021-9010.81.5.557>.
- Wall, J. E. (2000). Technology-delivered assessment: Power, problems, and promise. *Cybercounseling and Cyberlearning: Strategies and Resources for the Millennium*, 237–251. American Counseling Association.
- Wegge, J. (2006). Communication via videoconference: Emotional and cognitive consequences of affective personality dispositions, seeing one's own picture, and disturbing events. *Human-Computer Interaction*, 21(3), 273–318. https://doi-org.ezproxy.utu.fi/10.1207/s15327051hci2103_1.

- Woehr, D. J., & Arthur, W., Jr. (2003). The construct-related validity of assessment center ratings: A review and meta-analysis of the role of methodological factors. *Journal of Management*, 29(2), 231–258. <https://doi-org.ezproxy.utu.fi/10.1177/014920630302900206>.
- Youyou, W., Kosinski, M., & Stillwell, D. (2015). Computer-based personality judgments are more accurate than those made by humans. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, 112(4), 1036–1040. <https://doi-org.ezproxy.utu.fi/10.1073/pnas.1418680112>.
- Zygouris, S., & Tsolaki, M. (2015). Computerized cognitive testing for older adults: A review. *American Journal of Alzheimer's Disease and Other Dementias*, 30(1), 13–28. <https://doi.org/10.1177/1533317514522852>.